

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:57:56 ; Search time 27.86 Seconds
(without alignments)
227.059 Million cell updates/sec

Title: US-09-322-732-1

Perfect score: 946

Sequence: 1 MISVNDPKTGLTISVDNAIW.....VNEGDVLIINTGDSYISRK 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	946	100.0	185	21 Y70152	Staphylococcus aur
2	383.5	40.5	190	20 Y37395	Protein involved i
3	363.5	38.4	190	20 Y35550	Chlamydia pneumoni
4	280	29.6	185	20 Y37084	Protein involved i
5	270	28.5	185	20 Y34781	Chlamydia pneumoni
6	88.5	9.4	1174	11 R08257	B.thuringiensis to
7	87.5	9.2	2179	7 P60243	Sequence encoding
8	86.5	9.1	1174	13 R25825	Novel toxin expres
9	86.5	9.1	1174	17 R89494	B.t. toxin 81A2.
10	86.5	9.1	1174	18 W09043	Bacillus thuringie
11	84	8.9	2089	17 W08333	Cyclotella cryptic
12	83	8.8	436	20 Y35033	Chlamydia pneumoni

13	81	8.6	293	14	R35439	B. burgdorferi str
14	80.5	8.5	356	17	R95661	Immunogenic carrie
15	80.5	8.5	1511	20	Y32080	Yeast weak acid pu
16	78.5	8.3	349	17	R95660	Streptococcal prot
17	78	8.2	257	14	R36675	30 kD Borrelia bur
18	78	8.2	311	20	Y29854	Staphylococcus aur
19	78	8.2	841	20	W88420	Chlamydia pneumoni
20	78	8.2	2366	17	R95011	C. difficile toxin
21	78	8.2	2366	19	W68388	Clostridium diffic
22	77.5	8.2	425	8	P70315	Sequence of flagel
23	77.5	8.2	521	20	W96650	Fusion partner BB
24	77.5	8.2	736	20	Y24798	Recombinant strept
25	77	8.1	441	19	Y20612	Human neurofilamen
26	76.5	8.1	263	18	W55467	H. pylori ORF 05ee
27	76.5	8.1	300	18	W55571	H. pylori ORF 06ep
28	76.5	8.1	300	18	W55687	H. pylori ORF 06ep
29	76.5	8.1	443	20	Y19785	B. burgdorferi ant
30	76.5	8.1	447	8	P70314	Sequence of flagel
31	76.5	8.1	3623	20	Y27020	Rat cubilin protei
32	76	8.0	202	20	Y34607	Chlamydia pneumoni
33	76	8.0	430	20	Y19786	B. burgdorferi ant
34	76	8.0	663	17	R88732	S. aureus topoisome
35	76	8.0	771	16	R80495	Japanese oyster tr
36	76	8.0	1254	11	R07503	Meroprote apical-e
37	75.5	8.0	1254	18	W24575	CryET4. Bacillus
38	75.5	8.0	1167	15	R54073	Meroprote apical-e
39	75.5	8.0	1167	18	R35258	Bacillus thuringie
40	75.5	8.0	1167	18	W17700	CryET4. Bacillus
41	75.5	8.0	1167	20	Y30922	B. thuringiensis c
42	75.5	8.0	1167	20	W87632	CryET4 protein seq
43	74.5	7.9	142	18	W10208	Canola cold respon
44	74.5	7.9	183	19	W98704	H. pylori GHPO 707
45	74.5	7.9	214	18	W13629	Group G Streptococ

ALIGNMENTS

RESULT 1	
Y70152	
ID Y70152 standard; Protein; 185 AA.	
XX	
AC Y70152:	
XX	
DT 06-JUN-2000 (first entry)	
XX	
DE Staphylococcus aureus protein (HGS040).	
XX	
KW Vaccine; antibacterial; prevention; attenuation; detection;	
KW Staphylococcal infection; neonatal conjunctivitis; skin infection;	
KW toxic shock syndrome; osteomyelitis.	
XX	
OS Staphylococcus aureus.	
XX	
FH Key	Location/Qualifiers
FT Domain	30...32
FT Domain	/label= Antigenic_epitope
FT Domain	76..78
FT Domain	/label= Antigenic_epitope
XX	
PN WO200012678-A2.	
XX	
PD 09-MAR-2000.	
XX	
PF 31-AUG-1999; 99WO-US19726.	
XX	
PR 01-SEP-1998; 98US-0098964.	
XX	
PA (HUMA-) HUMAN GENOME SCI INC.	
XX	
PI Bailey CC, Choi GH;	
XX	
DR WPI; 2000-237864/20.	

Handwritten signature: C. J. J. J.

Sequence 190 AA;

Query Match 29.6%; Score 280; DB 20; Length 185;
Best Local Similarity 31.7%; Pred. No. 1.4e-20;
Matches 59: Conservative 44: Mismatches 79: Indels

[illegible]

RESULT 5
Y34781
ID Y34781 standard; Protein; 185 AA.

Protein involved in transcription, translation and/or maturation.

DE Chlamydia pneumoniae transmembrane protein sequence.

PN WO9927105-A2.

03-JUN-1999

XXXX

PF 20-NOV-1998; 98WO-IB01890.

04-NOV-1998: 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

XX

XX
XXXXXX (date)

PI Griffais R;

XX

DR WPI; 1999-357842/30.

Genome sequence of Chlamydia pneumoniae

XX
11
genome sequence of *Cytomydia pneumoniae*

PS Page 759-760; Disclosure; 1912pp; English.

```

XX Y34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 185 AA;

Query Match 28.5%; Score 270; DB 20; Length 185;
Best Local Similarity 31.2%; Pred. No. 1.4e-19;
Matches 58; Conservative 45; Mismatches 79; Indels 4; Gaps 2;

QY 1 MISVNDKFTGLTISVDNAIMKVIDFOHVKPKGSFAFVRSKLRNLRGTGAIOEKTRAGEKV 60
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 1 mvlssqlsvgmfstkdglkykvtsvkvgpkgsfikkvalqaadsdvliernfkateqv 60

QY 61 EPAMENRRMOYLYADGDNHVMFDMNESFEQTELSDDLKKEELNYLKEGMEVQIQTYEGET 120
   : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 61 keaqfetrileylyledesylfldlgnyeklifpqeikmndnflfkagvtvsamvydnv 120

QY 121 IGVELPKTVLTVTETPEGLKGD--TATGATKSATVETGYTLNVPFLVNECDVLIINTGD 178
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db 121 fsvelphflelmvsktd--fpgdlsisgvgkalletgievmvppiveigdvikidtrt 178

QY 179 GSIISR 184
   I I I
Db 179 ceylqr 184

RESULT 6
R08257
ID R08257 standard; protein; 1174 AA.
XX
AC R08257;
XX
DT 07-MAR-1991 (first entry)
XX
DE B.thuringiensis toxin gene product.
XX
KW Lepidopteran pests; insecticide.
XX
OS Bacillus thuringiensis.
XX
PN EP401979-A.
XX
PD 12-DEC-1990.
XX
PF 09-MAY-1990; 90EP-0304996.
XX
PR 14-DEC-1989; 89US-04511389.
XX
PR 18-MAY-1989; 89US-0353860.
XX
PA (MYCO-) MYCOGEN CORP.
XX
PI Payne J, Sick AJ;
XX
DR WPI; 1990-370000/50.
XX
DR N-PSDB; Q06830.
XX
DR New strains of Bacillus thuringiensis active against Lepidoptera
PT - and their new toxins and DNA encoding them
XX
PS Claim 3; Page 11-14; 20pp; English.
XX
CC Toxin product may be expressed from a transformed expression system

```

```

CC esp. E.coli and Pseudomonas fluorescens for insecticidal applications.
CC The toxins may be formulated with a phagostimulant or attractant for
CC beetles.
XX
SQ Sequence 1174 AA;

Query Match 9.4%; Score 88.5; DB 11; Length 1174;
Best Local Similarity 30.9%; Pred. No. 2.9;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;

QY 42 RNLRGTGAIOEKTRAGEKVPEPAMENRRMOYLYADGDNHVMFDMNESFEQTELSDDLKKEE 101
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db 1085 rnirayedayemmttasvnykptyeeer---ytdvggnhceydrvnyrppagvytke 1141

QY 102 LNYLKEGMEVQIQTYEGETIGVELPKTVLTVTTE 135
   I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db 1142 leyfpetdkvwiei--getegkfivdhvllme 1173

RESULT 7
P60243
ID P60243 standard; Protein; 2179 AA.
XX
AC P60243;
XX
DT 07-AUG-1991 (first entry)
XX
DE Sequence encoding the entire genomic RNA of human rhinovirus.
XX
KW Monoclonal antibody; MAB; HRV; vaccine; ss.
XX
OS Human rhinovirus.
XX
FH Key Location/Qualifiers
FT Region 1..69 /label= VP4 structural protein
FT Region 70..331 /label= VP2 structural protein
FT Region 332..567 /label= VP3 structural protein
FT Region 568..856 /label= VP1 structural protein
FT Region 857..1002 /label= 3B protein
FT Region 1003..1099 /label= 5B protein
FT Region 1100..1429 /label= X protein
FT Region 1430..1514 /label= protein 1B
FT Region 1515..1537 /label= protein VPg
FT Region 1538..1719 /label= Protease
FT Region 1720..2179 /label= Replicase
XX
PN EP169146-A.
XX
PD 22-JAN-1986.
XX
PF 17-JUL-1985; 85EP-0401465.
XX
PR 10-APR-1985; 85US-0721735.
XX
PR 20-JUL-1984; 84US-0632785.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Colonno RJ, Mitzutani S;
XX
PS WPI; 1986-022809/04.
XX
DR N-PSDB; N60194.

```



```

XX New DNA encoding the entire genomic RNA of human rhinovirus 14 -
PT and monoclonal antibodies which block attachment or neutralise
PT infectivity of rhinovirus.
XX
XX Example 11; Page 22-39; 80pp; English.
XX
CC Sequence may be used for the manufacture of hybridoma cells
CC expressing the HRV or fragments thereof. The fusion products may be
CC used in immunisation, or to raise MAbs for passive treatment of HRV
CC infection.
XX
XX Sequence 2179 AA;
SQ
Query Match 9.2%; Score 87.5; DB 7; Length 2179;
Best Local Similarity 24.5%; Pred. No. 8.5;
Matches 39; Conservative 25; Mismatches 74; Indels 21; Gaps 7;
QY 26 QHVKPGKGSFVRSKLNLRTGALQETFRAGEKVEPAMIENRRMOYLYAGDDNH--VFM 83
Db :| | | | : | | : | | : | | : | | : | |
1713 qyfvkdgqgviarhkvrfeinpvntat---ksklhpsvf-----ydvfpdgkpeavls 1763
QY 84 DNSEFQETLSSDYLKRELNVLK---EGMEVQIQTVGEGETIGVELPKTVELTVTETEPG 139
Db :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
1764 dndprlvkittesllsxykgnvnteptenmlvaydhagqllsidip-tseitikealyg 1822
QY 140 IKGDTATGATKSAFVETGYTLNVPLFVNEGDLVINTGD 178
Db :| | | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
1823 vdglepiddtsa---gfpv-vslgikkrdlmketqd 1856
RESULT 8
R25825
ID ID R25825 standard; Protein; 1174 AA.
XX AC R25825;
XX
DT 26-JAN-1993 (first entry)
XX
DE Novel toxin expressed by PS81A2.
XX
KW Bacillus thuringiensis; lepidoptera; insect pest; insecticide; ss.
XX
OS Bacillus thuringiensis.
XX
PN AU9056291-A.
XX
PD 18-JUN-1992.
XX
XX 05-JUN-1990; 90AU-0056291.
XX
XX 05-JUN-1990; 90AU-0056291.
XX
XX (MYCO ) MYCOGEN CORP.
XX
XX WPI; 1992-259364/32.
DR DR N-PSDB; Q26928.
XX
XX Controlling insect pests of lepidoptera family - includes
PT contacting insects with suitable amt. of Bacillus thuringiensis
XX
XX Claim 15; Page 29; 49pp; English.
XX
CC The protein sequence shows a Bacillus thuringiensis endotoxin whose
CC DNA may be used in an expression system to transform a variety of
CC microbial hosts, e.g. pseudomonas, Azobacter, Erwinia, Serratia,
CC Agrobacterium, Streptomyces species, etc. Expression of the toxic
CC gene results in the intracellular prodn. and maintenance of the
CC peptide toxin. With suitable hosts, the microbes can be applied to
CC the sites of lepidopteran insects, e.g. to the insects themselves,
CC to the rhizosphere, phylloplane or to a body of water, where they
CC will proliferate and be ingested by the insects. The result is the

```

CC	control of unwanted insects.. Alternatively, the microbe hosting the toxin gene can be treated under conditions that prolong the activity of the toxin produced in the cell. The treated cell can be applied to the environment of target pests. The resulting prod. retains the toxicity of the B.t. toxin. See also R25826.
XX	
SQ	Sequence 1174 AA;
 Query Match 9.1%; Score 86.5; DB 13; Length 1174; Best Local Similarity 30.9%; Pred. No. 4.6; Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps	
QY	42 RNLTGALGKTFRAGEKEVPEAMENRRNQYLADGDNHVFMDSFQTSLSSDYLKEE 101 : : : : :
Db	1085 rniryedayemnttasvnykpyeeer---ytdvgdghnceydrgvnrvpvagvytke 1141
QY	102 LNTLKGEVQIQTYEGETIGVELPKTVELTYTE 135 : :
Db	1142 lefypetdkwlei--getegkfivdnvellme 1173
 RESULT 9	
R89494	
ID	R89494 standard; Protein; 1174 AA.
XX	
AC	R89494;
XX	
DT	28-MAY-1996 (first entry)
XX	
DE	B.t. toxin 81A2.
XX	
KW	Toxin; delta-endotoxin; Bacillus thuringiensis; crystal protein;
XW	biological control; insecticide; pesticide; Lepidoptera; CryIF;
KW	Insect resistance; transgenic plant.
XX	
OS	Bacillus thuringiensis isolate PS81A2.
XX	
PN	WO9605314-A2.
XX	
PD	22-FEB-1996.
XX	
PF	14-AUG-1995; 95WO-US10310.
XX	
PR	15-AUG-1994; 94US-0291368.
XX	
PA	(MYCO) MYCOGEN CORP.
XX	
PI	Narva KE, Payne J, Schnepf HE, Schwab GE, Sick AJ;
XX	
DR	WFI: 1996-139708/14.
DR	N-PSDB; T16558.
XX	
PT	Toxin encoded by Bacillus thuringiensis PS91C2 nucleotide sequence -
PT	and isolates and delta endotoxins obtd. from it, used to control
PT	lepidopteran pests.
XX	
PS	Disclosure; Page 28-31; 52pp; English.
XX	
CC	Toxin 81A2 (R89494) is active against lepidopteran (caterpillar)
CC	pests. It is the product of a gene (T16558) isolated from Bacillus
CC	thuringiensis isolate PS81A2. B.t. isolates expressing CryIF class
CC	toxin genes can be used for the biological control of lepidopteran
CC	pests. Expression of the toxin by transgenic plants improves
CC	resistance to such insects.
XX	
SQ	Sequence 1174 AA;
 Query Match 9.1%; Score 86.5; DB 17; Length 1174; Best Local Similarity 30.9%; Pred. No. 4.6; Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps	

Query Match 9.1%; Score 86.5; DB 17; Length 1174;
Best Local Similarity 30.9%; Pred. No. 4.6;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;

	85	NSEFQETELSSDYLKEEENLYLKEG-MEVQIQTVTEGTIGVEL-PK-TVEITVTTETEPGIK	141
Df	380	kettfhemelaacrltqnlggydgagtveylynaadnkkfflelnprlqvelhpvt---gi-	435
QY	142	GDTATCATKSATVTGYTTLNPLF	165
Df	436	-----tganelpat-qiqvangiplf	454
		: : :	
RESULT	12		
ID	Y35033	standard; Protein; 436 AA.	
AC	XX		
CC	Y35033;		
DD	13-SEP-1999	(first entry)	
DE	Chlamydia pneumoniae transmembrane protein sequence.		
DX			
KW	Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;		
KW	sinusitis; purulent otitis media; erythema nodosum; pharyngitis;		
KW	vaccine; neutralising epitope.		
OS	Chlamydia pneumoniae.		
PX	WO9927105-A2.		
PN			
PP	03-JUN-1999.		
PF			
PI	20-NOV-1998; 98WO-IB01890.		
PR	04-NOV-1998; 98US-O107078.		
PR	21-NOV-1997; 97FR-0014673.		
XX	{GEST } GENSET.		
PA			
PT	Griffais R;		
DR	WPI; 1999-357842/30.		
XX	Genome sequence of Chlamydia pneumoniae		
PS	Page 926-927; Disclosure; 1912pp; English.		
XX			
CC	X34584-Y35879 represent the proteins encoded by all the open reading frames in the complete genome (see X91990) of Chlamydia pneumoniae.		
CC	C. pneumoniae causes respiratory disease such as pneumonia and C. bronchitis and is thought to be a contributing factor in heart disease, sarcoidosis, sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The polypeptides encoded by the open reading frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in immunogenic compositions as vaccines. Vectors containing C. pneumoniae nucleotide sequences can also be used as immunogenetic compositions, especially where the vector directs the expression of a neutralising epitope of C. pneumoniae.		
SQ	Sequence 436 AA;		
	Query Match 8.8%; Score 83; DB 20; Length 436;		
	Best Local Similarity 24.3%; Pred. No. 2.6;		
	Matches 36; Conservative 23; Mismatches 49; Indels 40; Gaps		
QY	24 DFOHVKFG---KGSAFVRSKLNLRFTGA -	-----OQKTFFRA-----GE 58 :	
Df	185 dfghlkpalrkxisiftrkeilarperilqlgvpksgsewllsagknsalpgtsdqg	244 :	
QY	59 KVEPAMERNRMOYLADGDNDHVMDNESPEQTLSLDYLKEEL---	NVLKGMMEVQIQ 114 ::: :	
Df	245 qlakvllenevid-lvs-----qdacteytiivnsfsfeekvlpnyrevlkrdiasql	296 	
QY	115 TYEGETIGVELPKTWELTAIVTFEPGIK	142	

61 SAEGADVDMNLTYLYNDGEFWHFMMNNETFEQLSADAKAIGDNAKWLLDOAECIVTLWNG 120 QV

```

Db      61 SAEGADVDDMLTLYNDGEFWHFMNNETPEQLSADAKAIGDGNKALLDQACIVTTLWNG 120
Qy      121 QPISVTPPNFVELEIVDTDPGLKGDTAGTGKGPATLSTGAVVKVPFLFVQIGEVIKVDTRS 180
        |||||||
Db      121 QPISVTPPNFVELEIVDTDPGLKGDTAGTGKGPATLSTGAVVKVPFLFVQIGEVIKVDTRS 180
Qy      181 GEYYSRVK 188
        |||||||
Db      181 GEYYSRVK 188

RESULT    2
F82047
C:Species: Vibrio cholerae
C:translation elongation factor EF-P VC2660 [imported] - Vibrio cholerae (group O1 strain
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: F82047
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, E.
l., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MJID:20406833
A:Accession: F82047
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <HEI>
A:Cross-references: GB:AE003852; NID:g9657252; PIDN:AAF95801.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N1696i; biotype El Tor
C:Genetics:
A:Gene: VC2660
A:Map position: 1
C:Superfamily: translation elongation factor EF-P

```

```

Query Match          76.9%; Score 750; DB 2; Length 188;
Best Local Similarity 73.9%; Pred. No. 3.4e-60;
Matches 139; Conservative 26; Mismatches 23; Indels 0; Gaps 0;

QY      1 MATYSNDRAGLKIMLDGEPYAVEASEFVPGKGQAFARVKLRLLTGTREKTEKSTD 60
      ||| :||: ||||||| || :||:|||||| ||:||:|||| | |||||||
Db       1 MATVTNEFGGLKIMLDNPECVILENEYKPKGQAFNRVRINKLTKGVLEKTFKSGD 60
      ||| |||||: ||||||| ||||| ||||| ||||| ||||| ||||| |||||

QY     61 SAEGADVVDNLITLYLNDGEFWHFMMNNETEFEQSADAKAIGDNAKWLLDQAECIVTLWNG 120
      :|| |||||: ||||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 TAEVADVVDIDDLYLNDGEFYHFMNNSTEFQLAAADAKAVGENAKWLVENNTCMLTWNG 120
      ||| |||||: ||||||| ||||| ||||| ||||| ||||| ||||| |||||

QY     121 OPTSVTPNPVELEIVDTDCLGKDGTACTGGKPATLTSTGAVKVPLEVOICEVIKVDTRS 180
      ||| |||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      121 NPVATVPNPVELEVETDTPGVKGDTOGTGGKPATLTSTGAVRVPLFVQIGEVIKVDTRS 180
      ||| |||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY     181 GEYVSRYK 188
      ||| |||
Db      181 AEYVGSRVK 188
      ||| |||

RESULT      3
I64061
translation elongation factor EF-P - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
C:Accession: I64061
A:Götschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J. D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: I64061
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-188 <TIGR>
```

A: Cross-references: GB:U32717; GB:I42023; NID:g1573283; PIDN:AAC21989.1; PID:g1573292
C: Superfamily: translation elongation factor EF-p
C: Keywords: protein biosynthesis

	Query Match	76.8%	Score 749;	DB 2;	Length 188;
	Best Local Similarity	75.0%;	Pred. No. 4.2e-67;		
	Matches 141;	Conservative 20;	Mismatches 27;	Indels 0;	Gaps 0;
Qy	1	MATYYSNDFRAGLKIMLDGEPYAYEASEFVKPGKGQAFARVKKLRLLLTGTTRVEKTFKSTD	60		
Db	1	MATYTSDFKPLGFMDQGPCVICIVENEFVKPGKGQAFTRIRKLISGVKLVLDVNEKSCT	60		
Qy	61	SAEAGADVDMNI ¹ TVLYNDGEFWHPMNNETFEOLSADAKAIGDNAKWILLDAQECIVTLWNG	120		
Db	61	SVEAADVMDLNL ¹ TSYKDDAFWYEMHPE ¹ TEFEQSADAKAVGDAEKWILLDQADCITVLWNG	120		
Qy	121	QPISVTPPNFVELEIVDTDPGLKGD ¹ TAGTGKGPATLSTGA ¹ VKVVPFLFVQIGEVIKVDTRS	180		
Db	121	APIVTTPNFVELEIVDTDPCLKGD ¹ TAGTGKGPATLSTGA ¹ VKVVPFLFVQIGEVIKVDTRS	180		
Qy	181	GEYVSrvk	188		
Db	181	GEYVSrvk	188		

RESULT 4
B82554 translation elongation factor EF-P XF2473 [imported] - Xylella fastidiosa (strain 9a5)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: B82554
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82554
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-193 <STM>
A:Cross-references: GB:AE004035; GB:AE003849; NID:g9107661; PIDN:AAF85271.1; GSPDB:GN
A:Experimental source: strain 9a5C
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Arões-Neto, E.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.N.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshukalo, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2473
C:Superfamily: translation elongation factor EF-P

		Query Match	65.4%;	Score 638;	DB 2;	Length 193;
		Best Local Similarity	63.3%;	Pred. No.	4e-50;	
		Matches 119;	Conservative 26;	Mismatches 43;	Indels 0;	Gaps
Qy	1	MATYYSNDFRAGLKMIDGSPYAVERASEFYVKPGKGAFARVKLRLLLTGTFRVETKFSTD	60			
		:::: :: : :::: :: :	:	: :::::::::::: :	:	: :::::
Dd	6	MASYGMNDVGNMKILYNAPAVITDTFYYVPKGQAFTVRKYRIKSGRVQEVTMKTSD	65			
Qy	61	SAEAGDVVDNNLYLYNDGEWFHFMNNETEQI LSADAKAIGDNNAKWILLDQACIVILWNG	120			
		: : :::: :: : :::: :: :	:	: :::::::::::: :	:	: :::::
Dd	66	TLEAADVVDDMOYL SDGSYWHEMFQMDETPEOVQADKNMGGAEKWLKBEOQCVAVLWNG	125			

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606

A:Accession: B72021

A:Molecule type: DNA

A:Residues: 1-190 <ARN>

A:Cross-references: GB:AE001363; GB:AE001363; NID:94377212; PIDN:AAAD19033.1; PID:g437721

A:Experimental source: strain CWL029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255

A:Accession: B81517

A:Molecule type: DNA

A:Residues: 1-190 <REA>

A:Cross-references: GB:AE002256; GB:AE002256; NID:g7189893; PIDN:AAF38751.1; PID:g7189898

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: efp-2; CP0971

C:Superfamily: translation elongation factor EF-P

Query Match 45.7%; Score 445.5; DB 2; Length 190;

Best Local Similarity 46.3%; Pred. No. 7.5e-33;

Matches 87; Conservative 42; Mismatches 58; Indels 1; Gaps 1;

QY 1 MATYNSNDFRAGIKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTVEKTFKSTD 60

Db 1 MVRVSTSEFRVGRUEIDGQYLLQNDVKPGKGQAFNRKVNFLTGRVRIETYSKGE 60

QY 61 SAEGADVDMNLTYLNDGEFHFMMNETFEQLSADAKAIGDNKAKWLLDQAEICIVTLWNG 120

Db 61 SVETADIVERSMRLLYTDQCATFMDDETPEQEVFWEKLENIQWLLLEDTIYTLVYNG 120

QY 121 QPISVTPNPFVEIVDTPGLKGDTA-GTGGKRPATLSTGAVKVPFLFVQIGEVKIVDTR 179

Db 121 DVAVPEPIEFMELSIETAPGVRGDTASGRVLPKPAVNTGAKIMVPIFIDEGELVKVDTR 180

QY 180 SGEVVSRRV 187

Db 181 TGSYSRRV 188

RESULT 9

G81401

translation elongation factor EF-P Cj0551 [similarity] - Campylobacter jejuni (strain NC

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000

C:Accession: G81401

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chiller

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barr

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A:Reference number: A81250; MUID:20150912

A:Accession: G81401

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-189 <PAR>

A:Cross-references: GB:AL139075; GB:AL139075; NID:g6967817; PIDN:CA875187.1; PID:g696801

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: efp; Cj0551

C:Superfamily: translation elongation factor EF-P

Query Match 44.9%; Score 438; DB 2; Length 189;

Best Local Similarity 45.7%; Pred. No. 3.5e-32;

Matches 85; Conservative 33; Mismatches 68; Indels 0; Gaps 0;

QY 1 MATYNSNDFRAGIKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTVEKTFKSTD 60

Db 1 MASYSMDLKKGLKLEIDGIPFIVEYQHVKPGKGPAPFRIKISFIDGKVLKTFHAGD 60

QY 61 SAEGADVDMNLTYLNDGEFHFMMNETFEQLSADAKAIGDNKAKWLLDQAEICIVTLWNG 120

Db 61 KCEAPNLEDKTMOYLDDGENCQFMDTQTYEQVAISDDDDVGEAKKWLMDGMWVDVLFHNG 120

QY 121 QPISVTPNPFVEIVDTPGLKGDTAGTGKRPATLSTGAVKVPFLFVQIGEVKIVDTR 180

Db 121 KAIGVEVPQVVELKIETAPNFKGDTOGSNKKPATLETGAVVQIPFHVLEGEVIRVDTR 180

QY 181 GEYYSR 186

Db 181 GEYIER 186

RESULT 10

T30278

translation elongation factor EF-P - Synecchococcus sp. (PCC 7942)

C:Species: Synecchococcus sp.

A:Variety: PCC 7942

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000

C:Accession: T30278

R:Phung, L.T.; Haselkorn, R.

submitted to the EMBL Data Library, May 1996

A:Description: Genes encoding biotin carboxyl carrier protein and elongation factor P

A:Reference number: Z20804

A:Accession: T30278

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-185 <PHU>

A:Cross-references: EMBL:U59235; NID:gl399825; PID:gl399829; PIDN:AAB82025.1

C:Genetics:

A:Note: efp

C:Superfamily: translation elongation factor EF-P

Query Match 43.7%; Score 426; DB 2; Length 185;

Best Local Similarity 48.6%; Pred. No. 4.1e-31;

Matches 88; Conservative 25; Mismatches 68; Indels 0; Gaps 0;

QY 6 SNDFRAGIKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTTRVEKTFKSTDSEGA 65

Db 4 SNDFRTGTTIHDGAVWRVVEFLHVKPGKSAFVRTKLKNAKTGNVVKTRFRAGETVPOA 63

QY 66 DVVDMNLTYLNDGEFHFMMNETFEQLSADAKAIGDNKAKWLLDQAEICIVTLWNGQPIVS 125

Db 64 VLEKSTLQYTKDGDDEFVDMETYESRLTAATIGRVKYLKRGMEANVTWNCQVIEV 123

QY 126 TPNFVLEIVDTPGLKGDTAGTGKRPATLSTGAVKVPFLFVQIGEVKIVDTRSGEYVS 185

Db 124 ELPSVYVLEIETDPGVKGTATGTGTPAKVETGAQVWVPLFISVGERIKIDTRNDSYLG 183

QY 186 R 186

Db 184 R 184

RESULT 11

S74333

translation elongation factor EF-P - Synecchocystis sp. (strain PCC 6803)

N:Alternate names: hypothetical protein slr0434

C:Species: Synecchocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S74333

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocys

s.

A:Reference number: S74322; MUID:97061201

A:Accession: S74333

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-187 <KAN>

RESULT 13
D71967
translation elongation factor EF-P - *Helicobacter pylori* (strain J99)
C;Species: *Helicobacter pylori*
A;Variety: strain J99

Qy 61 SAEGADVDMNLTYLYNDGEFFWFHFMNETFEQLSADAKAIGDNKAWLLDQAEIVTLWNG 120

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	702	74.2	185	2	A69620	translation elonga	
2	541	57.2	185	2	T30278	translation elonga	
3	535	56.6	187	2	S74333	translation elonga	
4	480	50.7	187	2	B70658	probable efp prote	
5	443	46.8	185	2	B75558	translation elonga	
6	435	46.0	192	2	E70418	translation elonga	
7	434	45.9	189	2	G81401	translation elonga	
8	421	44.5	185	2	B72212	translation elonga	
9	409	43.2	187	2	A64542	translation elonga	
10	406	42.9	187	2	D71967	translation elonga	
11	383.5	40.5	190	2	E71475	translation elonga	
12	379.5	40.1	190	2	B81738	probable translati	
13	372	39.3	193	2	B82554	translation elonga	
14	371	39.2	188	2	F82047	translation elonga	
15	363.5	38.4	190	2	B72021	translation elonga	
16	363	38.4	187	2	G71312	probable translati	
17	362	38.3	188	2	S34443	translation elonga	
18	343.5	36.3	188	2	G82908	translation elonga	
19	337	35.6	190	2	H64202	translation elonga	
20	331.5	35.0	188	2	B71678	translation elonga	
21	327	34.6	190	2	S73451	translation elonga	
22	311	32.9	188	2	I64061	translation elonga	
23	291	30.8	185	2	B81708	translation elonga	
24	290.5	30.7	188	2	F83290	translation elonga	
25	284	30.0	275	2	B64986	hypothetical 30.9	
26	275	29.1	185	2	F71553	probable translati	
27	273.5	28.9	199	2	D82229	translation elonga	
28	270	28.5	185	2	G72110	translation elonga	
29	269.5	28.5	192	2	F70126	translation elonga	

Query Match	56.6%	Score 535;	DB 2;	Length 187;
Best Local Similarity	53.8%;	Pred. No. 5.9e-37;		
Matches	99; Conservative	39; Mismatches	46; Indels	0; Gaps

QY	1	MISVDFKTLGTTISVDNAIHWIDFQHVKGSGSAFVRSKLRLNLTGAIOEKFTFRAGEKV	60
		: : : : :	

QY 1 MISVNDKFTGISVDNAIWKVIDFQHVKPGKGSFVRSKLRNLRTGAIQKTFRAGEKV 60

C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; MUID:99287316
 A:Accession: B72212
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-185 <ARN>
 A:Cross-references: GB:AE001815; GB:AE000512; NID:g4982341; PIDN:AAD36827.1; PID:g498
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TMI763
 C:Superfamily: translation elongation factor EF-P

Query Match 44.5%; Score 421; DB 2; Length 185;
 Best Local Similarity 44.6%; Pred. No. 1.3e-27;
 Matches 82; Conservative 35; Mismatches 67; Indels 0; Gaps 0;
 Qv 1 MISVNDKFTGLTISVDNAIKWKVIDFQHVKPGKGSFVRSKLRLNLTGAIOEKTFRAGEKV 60

Db 1 MIEVGLKGMFIYDGEIYVLEASKHFMRGSGSLINTKLKNYKTFGVREVPSPGEKV 60
Qy 61 EPAMENRRMOYLYADGNHVMFMDNESFEQTELSDDLKELNLYLKEGMEVOIQTYEGET 120
Db 61 QEAELSFKAQYLRDGHYFWMIDDDYEQVALSEEEIGDKYKYLVENMEVDLVFHHGTP 120
Qy 121 IGVLPKTVLTVTETEPGKIGDGTATGATKSATVETGYTLNVPFVNEGDVLIINTGDGS 180
Db 121 IGIELPTTVLTVTETEPGKIGDGTATGATKSATVETGYTLNVPFVNEGDVLIINTGDGS 180
Qy 181 YISR 184
Db 181 YXGR 184

RESULT 9
A64342
translation elongation factor EF-P - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 29-Sep-1999
C:Accession: A64342
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467
A:Accession: A64342
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-187 <TOM>
A:Cross-references: GB:AE000538; GB:AE000511; NID:g2313263; PIDN:AAD07247.1; PID:g231326
C:Superfamily: translation elongation factor EF-P

Query Match 43.2%; Score 409; DB 2; Length 187;
Best Local Similarity 40.4%; Pred. No. 1.3e-26;
Matches 74; Conservative 45; Mismatches 64; Indels 0; Gaps 0;

Qy 2 ISVNDFTGLTISVDNAIKWIDFQHVKPGKGSFAFVRSKLRLNLTGAIOEKTFRAGEKVE 61
Db 3 IGMSELKGLKIELGGVPYRIVEQHVKPGKGAFAFVRAKIKSFGLDGKVIKTFHAGDKCE 62
Qy 62 PAMENRRMOYLYADGNHVMFMDNESFEQTELSDDLKELNLYLKEGMEVOIQTYEGETI 121
Db 63 EPNLVEKTMQYLYHDGTYQFMDIESYEQIALNDSQVGEASKWMLDGMQVQLLHNDKAI 122
Qy 122 GVLPKTVLTVTETEPGKIGDGTATGATKSATVETGYTLNVPFVNEGDVLIINTGDGSY 181
Db 123 SVDVPQVVALKIVETAPNFKGDTSSASKKPKATLETGAVQVPPHVLGEETIKVNTETEY 182
Qy 182 ISR 184
Db 183 LEK 185

RESULT 10
D71967
translation elongation factor EF-P - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 17-Mar-2000
C:Accession: D71967
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.
A:Reference number: A71800; MUID:99120557
A:Accession: D71967
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-187 <ARN>
A:Cross-references: GB:AE001454; GB:AE001439; NID:g4154666; PIDN:AAD05735.1; PID:g415
A:Experimental source: strain J99
C:Genetics:
A:Gene: efp
C:Superfamily: translation elongation factor EF-P

Query Match 42.9%; Score 406; DB 2; Length 187;
Best Local Similarity 40.4%; Pred. No. 2.3e-26;
Matches 74; Conservative 44; Mismatches 65; Indels 0; Gaps 0;

Qy 2 ISVNDFTGLTISVDNAIKWIDFQHVKPGKGSFAFVRSKLRLNLTGAIOEKTFRAGEKVE 61
Db 3 IGMSELKGLKIELGGVPYRIVEQHVKPGKGAFAFVRAKIKSFGLDGKVIKTFHAGDKCE 62
Qy 62 PAMENRRMOYLYADGNHVMFMDNESFEQTELSDDLKELNLYLKEGMEVOIQTYEGETI 121
Db 63 EPNLVEKTMQYLYHDGTYQFMDIESYEQIALNDSQVGEASKWMLDGMQVQLLHNDKAI 122
Qy 122 GVLPKTVLTVTETEPGKIGDGTATGATKSATVETGYTLNVPFVNEGDVLIINTGDGSY 181
Db 123 SVDVPQVVALKIVETAPNFKGDTSSASKKPKATLETGAVQVPPHVLGEETIKVNTETEY 182
Qy 182 ISR 184
Db 183 LEK 185

RESULT 11
E71475
probable translation elongation factor EF-P - Chlamydia trachomatis (serotype D, strain 282, 754-759, 1998)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: E71475
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A:Reference number: A71570; MUID:9900809
A:Accession: E71475
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <ARN>
A:Cross-references: GB:AE001347; GB:AE001273; NID:g3329210; PIDN:AAC68347.1; PID:g332
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: efp-2
C:Superfamily: translation elongation factor EF-P

Query Match 40.5%; Score 383.5; DB 2; Length 190;
Best Local Similarity 41.3%; Pred. No. 1.6e-24;
Matches 76; Conservative 39; Mismatches 68; Indels 1; Gaps 1;

Qy 2 ISVNDFTGLTISVDNAIKWIDFQHVKPGKGSFAFVRSKLRLNLTGAIOEKTFRAGEKVE 61
Db 4 VSTSEFRVGLRVKIDGQPYVILQNDVFKPGKGAFAFVRAKIKSFGLDGKVIKTFKSGESIE 63
Qy 62 PAMENRRMOYLYADGNHVMFMDNESFEQTELSDDLKELNLYLKEGMEVOIQTYEGETI 121
Db 64 TADVRFOOMRLTYDDEGATFMDDETFOELIYFMDKLENVROWLLEDTITVLYNGDVI 123
Qy 122 GVLPKTVLTVTETEPGKIGDGTATG-ATKSATVETGYTLNVPFVNEGDVLIINTGDGS 180
Db 124 SVEPPFMELTIAETAPGVGRDASGRVLKPKATNTGAKIMVPIFIEEGEVKVDRTGS 183
Qy 181 YISR 184
Db 184 YESR 187

RESULT 12
B81738

translation elongation factor P TC0133 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
A:Accession: B81738
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255
A:Accession: B81738
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <TET>
A:Cross-references: GB:AE002280; GB:AE002160; NID:g7190162; PIDN:AAF39011.1; PID:g719016
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
C:Superfamily: translation elongation factor EF-P

Query Match 40.1%; Score 379.5; DB 2; Length 190;
Best Local Similarity 40.8%; Pred. No. 3.5e-24;
Matches 75; Conservative 40; Mismatches 68; Indels 1; Gaps 1;
Qy 2 ISVNDFKTLTISVDNAIKWKVIDFQHVKPGKSAFVRSKLNRLRTCAIQEKTFRAGEKVE 61
Db 4 VSTSEFRVGLRVEDIDQGPYVILONDFVKPGKQAFNRVIRKLNFLTVRIETKSGESIE 63
Qy 62 PAMENRRMOYLYADGDHNVFMNDSFEQTELSDDYLKEELNLYLKEGMEVQIQTEGETI 121
Db 64 TADVREQQRLLYTQEGATFMDDETFEQELIFWDLKLNIRKWLLEDIVYTLVRNGDVI 123
Qy 122 GVLEPKTVELTETEPGKIGDTATG-ATKSAVTETGYTLNPLFVNEGDVLIINTGDS 180
Db 124 SVEPIFELSAETAAPVGRGDTASGRVLKPAITNTGAKINVPFIEBGEVVKVDTGTS 183
Qy 181 YISR 184
Db 184 YESR 187

RESULT 13
B82554
translation elongation factor EF-P XF2473 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
A:Accession: B82554
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82554
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-193 <SIM>
A:Cross-references: GB:AE004055; GB:AE003849; NID:g9107661; PIDN:AAF85271.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Bioness, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H. as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigachado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak M.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tshubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovsky-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:

A:Gene: XF2473
C:Superfamily: translation elongation factor EF-P
Query Match 39.3%; Score 372; DB 2; Length 193;
Best Local Similarity 40.9%; Pred. No. 1.5e-23;
Matches 74; Conservative 40; Mismatches 67; Indels 0; Gaps 0;
Qy 4 VDNFKTGLTISVDNAIKWKVIDFQHVKPGKSAFVRSKLNRLRTCAIQEKTFRAGEKVEPA 63
Db 11 MNDVKNMGKILVNAEPAVITTEYVVKPGKQAFTRVKYRLIKSGRQVETMKSTDLEAA 70
Qy 64 MIENRRMOYLYADGDHNVFMNDSFEQTELSDDYLKEELNLYLKEGMEVQIQTEGETIGV 123
Db 71 DVVDTDMOYLYSDGEYWHFMQOETFEQVQADKNGMGAEKWLKGEQECVVTLWNGVPIGV 130
Qy 124 ELPKTVELTETEPGKIGDTATGATKSAVTETGYTLNPLFVNEGDVLIINTGDSGIS 183
Db 131 QPPNFVELKITETDPLGRDITSGGGKIPATLETGAVVRVPLFVNDQEVKIKVDTRSGEYVS 190
Qy 184 R 184
Db 191 R 191

RESULT 14
F82047
translation elongation factor EF-P VC3660 [imported] - Vibrio cholerae (group O1 strain
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
A:Accession: F82047
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R. chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833
A:Accession: F82047
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <HEI>
A:Cross-references: GB:AE004332; GB:AE003852; NID:g9657252; PIDN:AAF95801.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Map position: 1
C:Superfamily: translation elongation factor EF-P

Query Match 39.2%; Score 371; DB 2; Length 188;
Best Local Similarity 41.0%; Pred. No. 1.7e-23;
Matches 75; Conservative 35; Mismatches 73; Indels 0; Gaps 0;
Qy 2 ISVNDFKTLTISVDNAIKWKVIDFQHVKPGKSAFVRSKLNRLRTCAIQEKTFRAGEKVE 61
Db 4 VSTNEFKGLKLTMLDNEPCVILENVEYKPGKQAFNRVIRKLNLTGKLVLETKSGDTAE 63
Qy 62 PAMENRRMOYLYADGDHNVFMNDSFEQTELSDDYLKEELNLYLKEGMEVQIQTEGETI 121
Db 64 VADVVDIDLDLYNDGGEYHFMNNTSTFQLAADAKAVGENAKWLVENWTCMLTLWNGNPI 123
Qy 122 GVLEPKTVELTETEPGKIGDTATGATKSAVTETGYTLNPLFVNEGDVLIINTGDSGY 181
Db 124 AVTPNFVELEVETEDPGVKGDGTQGTGKIPATLSTGAVVRVPLFVQIGEVKVDTRSAEY 183
Qy 182 ISR 184
Db 184 VGR 186

RESULT 15
B72021
translation elongation factor p CP0971 [imported] - Chlamydia pneumoniae (strains
A:Gene: XF2473
C:Superfamily: translation elongation factor EF-P
Query Match 39.3%; Score 372; DB 2; Length 193;
Best Local Similarity 40.9%; Pred. No. 1.5e-23;
Matches 74; Conservative 40; Mismatches 67; Indels 0; Gaps 0;
Qy 4 VDNFKTGLTISVDNAIKWKVIDFQHVKPGKSAFVRSKLNRLRTCAIQEKTFRAGEKVEPA 63
Db 11 MNDVKNMGKILVNAEPAVITTEYVVKPGKQAFTRVKYRLIKSGRQVETMKSTDLEAA 70
Qy 64 MIENRRMOYLYADGDHNVFMNDSFEQTELSDDYLKEELNLYLKEGMEVQIQTEGETIGV 123
Db 71 DVVDTDMOYLYSDGEYWHFMQOETFEQVQADKNGMGAEKWLKGEQECVVTLWNGVPIGV 130
Qy 124 ELPKTVELTETEPGKIGDTATGATKSAVTETGYTLNPLFVNEGDVLIINTGDSGIS 183
Db 131 QPPNFVELKITETDPLGRDITSGGGKIPATLETGAVVRVPLFVNDQEVKIKVDTRSGEYVS 190
Qy 184 R 184
Db 191 R 191
RESULT 14
F82047
translation elongation factor EF-P VC3660 [imported] - Vibrio cholerae (group O1 strain
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
A:Accession: F82047
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R. chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833
A:Accession: F82047
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <HEI>
A:Cross-references: GB:AE004332; GB:AE003852; NID:g9657252; PIDN:AAF95801.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Map position: 1
C:Superfamily: translation elongation factor EF-P
Query Match 39.2%; Score 371; DB 2; Length 188;
Best Local Similarity 41.0%; Pred. No. 1.7e-23;
Matches 75; Conservative 35; Mismatches 73; Indels 0; Gaps 0;
Qy 2 ISVNDFKTLTISVDNAIKWKVIDFQHVKPGKSAFVRSKLNRLRTCAIQEKTFRAGEKVE 61
Db 4 VSTNEFKGLKLTMLDNEPCVILENVEYKPGKQAFNRVIRKLNLTGKLVLETKSGDTAE 63
Qy 62 PAMENRRMOYLYADGDHNVFMNDSFEQTELSDDYLKEELNLYLKEGMEVQIQTEGETI 121
Db 64 VADVVDIDLDLYNDGGEYHFMNNTSTFQLAADAKAVGENAKWLVENWTCMLTLWNGNPI 123
Qy 122 GVLEPKTVELTETEPGKIGDTATGATKSAVTETGYTLNPLFVNEGDVLIINTGDSGY 181
Db 124 AVTPNFVELEVETEDPGVKGDGTQGTGKIPATLSTGAVVRVPLFVQIGEVKVDTRSAEY 183
Qy 182 ISR 184
Db 184 VGR 186
RESULT 15
B72021
translation elongation factor p CP0971 [imported] - Chlamydia pneumoniae (strains

Search completed: January 12, 2001, 14:59:23
Job time: 83 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:57:56 ; Search time 21.21 Seconds
(without alignments)
156.627 Million cell updates/sec

Title: US-09-322-732-1

Perfect score: 946

Sequence: 1 MISVNDFKTGLTISVDNAIW.....VNEGDVLIINTGDGVSIRG 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/1aa/6_COMB.pep.*

4: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*

5: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.5	9.1	1174	1	US-08-040-751-3
2	86.5	9.1	1174	1	US-08-291-368-2
3	86.5	9.1	1174	2	US-08-962-190-2
4	86.5	9.1	1174	4	PCT-US95-10310-2
5	86.5	9.1	1174	5	5164180-4
6	84	8.9	2089	1	US-08-418-893D-23
7	84	8.9	2089	1	US-08-418-893D-24
8	83	8.8	294	1	US-08-137-175A-9
9	83	8.8	294	3	US-08-479-017-9
10	80	8.5	1430	3	US-09-008-172-2
11	78	8.2	257	1	US-07-781-355-2
12	78	8.2	2366	1	US-08-480-604A-10
13	78	8.2	2366	2	US-08-405-496A-10
14	77.5	8.2	464	2	US-09-021-323-3
15	76	8.0	663	3	US-08-776-265-5
16	76	8.0	770	1	US-08-525-654A-1
17	76	8.0	771	1	US-08-525-654A-3
18	75.5	8.0	1167	1	US-08-100-709-2
19	75.5	8.0	1167	1	US-08-176-865-2
20	75.5	8.0	1167	1	US-08-474-038-2
21	75.5	8.0	1167	2	US-08-779-046-2
22	75.5	8.0	1167	2	US-08-881-340-2
23	74.5	7.9	801	3	US-09-104-070-2
24	74.5	7.9	1176	1	US-08-434-823-2
25	74.5	7.9	1176	1	US-08-457-366-2
26	74.5	7.9	1535	3	US-08-755-587-185
27	74	7.8	693	1	US-08-553-279-2
28	73	7.7	526	2	US-08-853-659A-40

29	73	7.7	999	2	US-08-770-301A-1	Sequence 1, Appli
30	73	7.7	999	3	US-09-175-581-1	Sequence 1, Appli
31	72	7.6	644	1	US-08-487-890A-6	Sequence 6, Appli
32	72	7.6	644	2	US-08-478-435-6	Sequence 6, Appli
33	72	7.6	644	2	US-08-337-483-6	Sequence 6, Appli
34	72	7.6	644	2	US-08-478-373-6	Sequence 6, Appli
35	72	7.6	644	3	US-08-474-671-6	Sequence 6, Appli
36	72	7.6	644	3	US-08-483-577A-6	Sequence 6, Appli
37	72	7.6	664	3	US-08-669-408B-2	Sequence 2, Appli
38	72	7.6	1683	3	US-08-755-587-183	Sequence 183, App
39	71	7.5	277	1	US-08-400-413-1	Sequence 1, Appli
40	71	7.5	348	2	US-08-844-153-2	Sequence 2, Appli
41	71	7.5	419	3	US-09-155-200-4	Sequence 2, Appli
42	71	7.5	480	2	US-08-962-203-2	Sequence 2, Appli
43	71	7.5	480	3	US-09-282-125A-2	Sequence 2, Appli
44	71	7.5	500	1	US-08-260-582-77	Sequence 77, Appli
45	71	7.5	500	4	PCT-US95-03471-77	Sequence 77, Appli

ALIGNMENTS

RESULT 1
US-08-040-751-3
; Sequence 3, Application US/08040751
; Patent No. 5407825
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M
; APPLICANT: Sick, August J
; TITLE OF INVENTION: No. 5407825el Bacillus thuringiensis isolates
; TITLE OF INVENTION: active against Lepidopteran Pests and Genes Encoding No. 54
; TITLE OF INVENTION: Lepidopteran-active Toxins
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID R. SALIWANCHIK
; STREET: 2421 N.W. 41st STREET, SUITE A-1
; CITY: GAINESVILLE
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/040,751
; APPLICATION NUMBER: 19930329
; FILING DATE: 19930329
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SALIWANCHIK, DAVID R.
; REFERENCE/DOCKET NUMBER: MA39.C1.D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1174 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem - 11 (tm) Library of August Sick
; CLONE: 81A2
; US-08-040-751-3

```
Query Match      9.1%; Score 86.5; DB 1; Length 1174;
Best Local Similarity 30.9%; Pred. NO. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps

QY   42 RNLRGTAIQEKTPRAGEKVEPAMINRRMQYLIYAGDGNHVMFNDSFEQTLESSDYLKEE 101
      ||| | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    1085 RNRIDAYEMNTASVNYKPTYESER---YTDVGDNHCHEVDRCGVNRPVPGYVTKE 1141

QY   102 LNVLKCEMEVOIQTYEGETIGVELPKTVELTVTE 135
      | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    1142 LEYFPPTDKWIEI--GETEGKFIVDNVELLINE 1173
```

```

RESULT      2
US-08-291-368-2
;
; Sequence 2, Application US/08291368
; Patent No. 5686069
;
; GENERAL INFORMATION:
;
; APPLICANT: Payne, Jewel M.
; APPLICANT: Sick, August J.
; TITLE OF INVENTION: No. 5686069el Bacillus thuringiensis
;                   Active Against Lepidopteran P
; NUMBER OF SEQUENCES: 27
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/291.368

```

FILING DATE: 536
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/597,607
 FILING DATE: 15-OCT-90
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Saliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCE/DOCKET NUMBER: MA50.C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (904)375-8100
 TELEFAX: (904)372-5800
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1174 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Bacillus thuringiensis
 STRAIN: aizawai
 INDIVIDUAL ISOLATE: PS81A2
 IMMEDIATE SOURCE:
 LIBRARY: Lambdaagem - 11 (tm) Lib
 CLONE: 81A2

US-08-291-368-2

Query Match 9.1%; Score 86.5; DB 1; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;

	Matches	29; Conservative	10; Mismatches	50; Indels	5; Gaps
Qy	42	NRLTGAIQEKTFRAGEKVEPAMLENRRMQLVYADGDNHVPMDNESFQTELTSSDYKKEE	101		
				:	:
Db	1085	NNIRYDAYENMTTASVNYKPTYEER---	YTDVGDNHCHEYDRGVNRYRVPAGYVTKE	1141	
Qy	102	LNLYLKEGMEVQIQTYEGETIGVLVLPKTVELTVTE	135		
Db	1142	LEYEPETDKVMVIEI--GETEGKFTVDNVNELLME	1173		

RESULT 3
 US-08-962-190-2
 ; Sequence 2, Application US/08962190
 ; Patent No. 5985267
 ; GENERAL INFORMATION:
 ; APPLICANT: Payne, Jewel M.
 ; APPLICANT: Sick, August J.
 ; TITLE OF INVENTION: No. 5985267el Bacillus thuringiensis Isolates
 ; TITLE OF INVENTION: Active Against Lepidopteran Pests
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Saliwanchik & Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: FL
 ; COUNTRY: US
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/962,190
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/291,368
 ; FILING DATE:
 ; APPLICATION NUMBER: 07/597,607

```

1  FILING DATE: 15-OCT-90
2  ATTORNEY/AGENT INFORMATION:
3  NAME: Saliwanchik, David R.
4  REGISTRATION NUMBER: 31,794
5  REFERENCE/DOCKET NUMBER: MA50.C1
6  TELECOMMUNICATION INFORMATION:
7  TELEPHONE: (904)375-8100
8  TELEFAX: (904)372-5800
9  INFORMATION FOR SEQ ID NO: 2:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 1174 amino acids
12 TYPE: amino acid
13 STRANDEDNESS: single
14 TOPOLOGY: linear
15 MOLECULE TYPE: protein
16 HYPOTHEetical: YES
17 ANTI-SENSE: NO
18 ORIGINAL SOURCE:
19 ORGANISM: Bacillus thuringiensis
20 STRAIN: aizawai
21 INDIVIDUAL ISOLATE: PS81A2
22 IMMEDIATE SOURCE:
23 LIBRARY: Lambdaagem - 11 (tm) Lib
24 CLONE: 81A2
25 US-08-962-190-2

```

Query Match 9.1%; Score 86.5; DB 2; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;

QY 42 RNLRTGAIQEKTFRAGEKVEPAMIENRRMQYLYADGDNHVMDNESFEQTELSDDLKEE 101

Db 1085 RNIRYEDAYEMNTTASVNYKPTYEEER---YTDVQGDHNCHEYDRGYVNYRVPVAGYVTKE 1141
Qy 102 LNYLKEGMEVOIQTYEGTIGVELPKTVELTVTE 135
Db 1142 LEYFPETDKVWIEI--GETEGKFIVDNVELLME 1173

RESULT 4
PCT-US95-10310-2
; Sequence 2, Application PC/TUS9510310
; GENERAL INFORMATION:
; APPLICANT: MYCOGEN CORPORATION
; APPLICANT: STREET ADDRESS: 5501 Oberlin Drive
; APPLICANT: CITY: San Diego
; APPLICANT: STATE/PROVINCE: California
; APPLICANT: COUNTRY: US
; APPLICANT: POSTAL CODE/ZIP: 92121
; APPLICANT: PHONE NUMBER: (619) 453-8030
; APPLICANT: FAX NUMBER: (619) 453-6991
; TITLE OF INVENTION: Protein Toxins Active Against Lepidopteran Pests
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10310
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/597,607
; FILING DATE: 15-OCT-90
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA50.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904)375-8100
; TELEFAX: (904)372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1174 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem - 11 (tm) Library of August Sick
; CLONE: 81A2
PCT-US95-10310-2

Query Match 9.1%; Score 86.5; DB 4; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;
Qy 42 RNLRTGATQKTFRAGEKVEPAMENRRMQYLYADGDNHVFMDNESFQTELSSDYLKEE 101

Db 1085 RNIRYEDAYEMNTTASVNYKPTYEEER---YTDVQGDHNCHEYDRGYVNYRVPVAGYVTKE 1141
Qy 102 LNYLKEGMEVOIQTYEGTIGVELPKTVELTVTE 135
Db 1142 LEYFPETDKVWIEI--GETEGKFIVDNVELLME 1173

RESULT 5
5164180-4
; Patent No. 5164180
; APPLICANT: Payne, Jewel; Sick, August J.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS ISOLATES ACTIVE
; AGAINST LEPIDOPTERAN PESTS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/451,389
; FILING DATE: 14-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 353,860
; FILING DATE: 18-MAY-1989
; SEQ ID NO:4:
; LENGTH: 1174
5164180-4

Query Match 9.1%; Score 86.5; DB 5; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;
Qy 42 RNLRTGATQKTFRAGEKVEPAMENRRMQYLYADGDNHVFMDNESFQTELSSDYLKEE 101
Db 1085 RNIRYEDAYEMNTTASVNYKPTYEEER---YTDVQGDHNCHEYDRGYVNYRVPVAGYVTKE 1141
Qy 102 LNYLKEGMEVOIQTYEGTIGVELPKTVELTVTE 135
Db 1142 LEYFPETDKVWIEI--GETEGKFIVDNVELLME 1173

RESULT 6
US-08-418-893D-23
; Sequence 23, Application US/08418893D
; Patent No. 5559220
; GENERAL INFORMATION:
; APPLICANT: ROESSLER, PAUL G
; APPLICANT: OHLROGGE, JOHN B
; TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A
; TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTHELLA CRYPTICA
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY
; STREET: 1617 Cole Blvd.
; CITY: Golden
; STATE: CO
; COUNTRY: USA
; ZIP: 80401-3393
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,893D
; FILING DATE: April 7, 1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,938
; FILING DATE: September 14, 1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: O'CONNOR, EDNA
; REGISTRATION NUMBER: 29,252
; REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON

RESULT 7
US-08-418-893d-24
; Sequence 24, Application US/08418893D
; Patent No. 5559220
; GENERAL INFORMATION:
; APPLICANT: ROESSLER, PAUL G
; APPLICANT: OHROEGGE, JOHN B
; TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A
; TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTELLA CRYPTICA
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY
; STREET: 1617 Cole Blvd.
; CITY: Golden
; STATE: CO
; COUNTRY: USA
; ZIP: 80401-3393
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,893D
; FILING DATE: April 7, 1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,938
; FILING DATE: September 14, 1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: O'CONNOR, EDNA
; REGISTRATION NUMBER: 29,252
; REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-231-1000

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,175A
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08972
FILING DATE: 22-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: BARBOUR-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528

```
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-137-175A-9

Query Match      8.8%; Score 83; DB 1; Length 294;
Best Local Similarity 24.08; Pred. No. 0.32;
Matches 47; Conservative 30; Mismatches 61; Indels 58; Gaps 11;

Qy 1 MISVNDKTKTGLTISVDNAIKWVIDFOHVKPGKGSFVRSKLRNLTGAIOEKTFRAGEKV 60
Db 110 MLVSDLLNT-ITIEYDPSNKKISSQVAK-----KQGSLTEETTKT-SKL 152
Qy 61 EPAMIEARMQVLYADGNHVFMDNESFEQTELS-SDYLKEELNLYLKEGMEVQIOTVEGE 119
Db 153 SAKKITR-----SNNITIEYTEMTDADNASKAVETLNGI-----TLEGS 192
Qy 120 TIGVELPKTV-ELFVTV-----ETEPGK---GDTATGATKSATV--ETGYTLNVP----- 163
Db 193 LVGGKTTLTITKEGTVTLKKEIEKAGTVKLFDDLTASSATKKTAVWNDTSSTLTVSAGEKK 252
Qy 164 ----LFVNEGDVLIIN 175
Db 253 TKDFVFLTDGTTIVQN 268

RESULT 9
US-08-479-017-9
; Sequence 9, Application US/08479017
; Patent No. 6143872
; GENERAL INFORMATION:
; APPLICANT: BARBOUR, Alan G.
; APPLICANT: BERGSTROM, Sven
; APPLICANT: HANSSON, Leenart
; TITLE OF INVENTION: IMPROVEMENT IN BORRELIA BURGDORFERI AND
; TITLE OF INVENTION: PROPHYLAXIS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,017
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,175
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: PCT/US92/08972
; FILING DATE: 22-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: BARBOUR-1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-479-017-9

Query Match      8.8%; Score 83; DB 3; Length 294;
Best Local Similarity 24.08; Pred. No. 0.32;
Matches 47; Conservative 30; Mismatches 61; Indels 58; Gaps 11;

Qy 1 MISVNDKTKTGLTISVDNAIKWVIDFOHVKPGKGSFVRSKLRNLTGAIOEKTFRAGEKV 60
Db 110 MLVSDLLNT-ITIEYDPSNKKISSQVAK-----KQGSLTEETTKT-SKL 152
Qy 61 EPAMIEARMQVLYADGNHVFMDNESFEQTELS-SDYLKEELNLYLKEGMEVQIOTVEGE 119
Db 153 SAKKITR-----SNNITIEYTEMTDADNASKAVETLNGI-----TLEGS 192
Qy 120 TIGVELPKTV-ELFVTV-----ETEPGK---GDTATGATKSATV--ETGYTLNVP----- 163
Db 193 LVGGKTTLTITKEGTVTLKKEIEKAGTVKLFDDLTASSATKKTAVWNDTSSTLTVSAGEKK 252
Qy 164 ----LFVNEGDVLIIN 175
Db 253 TKDFVFLTDGTTIVQN 268

RESULT 10
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nicholls, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-008-172-2

Query Match      8.5%; Score 80; DB 3; Length 1430;
Best Local Similarity 22.2%; Pred. No. 7;
Matches 40; Conservative 31; Mismatches 87; Indels 22; Gaps 6;

Qy 15 VDNAIKWVIDFOHVKPGKGSFVRSKLRNLTGAIOEKTFRAGEK-----VEPAMI 65
Db 587 VQTVIAKIKAQINPKTDGLTFLDELK--QAFKIYNEDMRQAKKKYQTSNIPTAYALML 644
Qy 66 ENR----RMQY--LYADGNHVFMDNESFEQTELSDDLKEELNLYLKEGMEVQIOTVEGE 119
Db 645 SNKDSITRLYYGDMYSDGQYMATKSPYYDAID---TLLKARIKYAAGQDMKITVYVEGD 701
Qy 120 TIGVELPKTVELTETEPGKIGKDTATG--ATKSATVETGYTLNVPFVNEGDVLIINTG 177
Db 702 KSHMDWDYTGVLTSVRYGTGANEATDQGEATKTQGMAVITSNPNPSLKLNDKDVIVNNG 761

RESULT 11
US-07-781-355-2
; Sequence 2, Application US/07781355
; Patent No. 5246844
; GENERAL INFORMATION:
; APPLICANT: No. 5246844ris, Steven J.
```

APPLICANT: Barbour, Alan G.
TITLE OF INVENTION: VIRULENCE ASSOCIATED PROTEINS IN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/781,355
FILING DATE: 19911022
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: USH:162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-474-7577
TELEFAX: 512-474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-781-355-2

Query Match 8.2%; Score 78; DB 1; Length 257;
Best Local Similarity 24.2%; Pred. No. 0.92;
Matches 39; Conservative 24; Mismatches 64; Indels 34; Gaps 7;

Qy 23 IDQHVKPGKSAFVRSKLNRLRTGAIOEKTFRAGEKVEPAMENRRMOYLYADGDNHVF 82
Db 124 IDLEKIKNSDKAIFASML-----AKEAYSLTKAAEONMOKLY----- 161
Qy 83 MDNESFEQTESSDYL-KEELNYLKEGMEVQIQ-TYEGETIGVELPKTVLTVTETEPGI 140
Db 162 KEQOKISESESDYSDSAEIKQAKEAVEIAWKAATVEAKDKLID----VENTVKTLDKI 217
Qy 141 KGDATGATKATVETGTYTLNVPFVNGDVL-----IINT 176
Db 218 KTET-TNNTKLADIKEAAELVLQIAKNAKEIVQEVVALLNT 257

RESULT 12
US-08-480-604A-10
Sequence 10, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2366 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-604A-10

Query Match 8.2%; Score 78; DB 1; Length 2366;
Best Local Similarity 21.5%; Pred. No. 24;
Matches 50; Conservative 39; Mismatches 74; Indels 70; Gaps 13;

Qy 1 MISVDPK-----TGLTISVDNAIWKVI---DFQHVKPGKGSFVRSK--- 40
Db 736 IVSANQYEVNRINSEGRRELLDHSGEWINKESIIKDISKEYISFNPKENKITVKSNNLP 795
Qy 41 -----LRNLRGTGA-----IOEKT-----FRAGEKVEPAMENR---RMQYLYADGDNHV 81
Db 796 ELSTLLQEIENNNSSSDIELEEKVMLTCEINVISNIDTQIVEEIRIEEAKNLTSDSINYI 855
Qy 82 FMDNESFEQTESSDYL-----KEELNYLKEGMEVQ-IQTYEGETIGVE 124
Db 856 ---KDEPKLIESDLCDLKQONELEDHSFISFEDSETDEGSIRFINKEGESIFVE 912
Qy 125 LPKTV-----ELTVTETEPGKIG---DPATG-ATKSATVETGY---TLNVPFLV 166
Db 913 TEKTIFSEYANHTELSKIKGTIFDTVNGKLVKKNLDTTHEVNTLNAAFFI 965

RESULT 13
US-08-405-496A-10
Sequence 10, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
NEUROTOXIN
NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,496A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPHD-01308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-405-496A-10

Query Match 8.2%; Score 78; DB 2; Length 2366;
Best Local Similarity 21.5%; Pred. No. 24;
Matches 50; Conservative 39; Mismatches 74; Indels 70; Gaps 13;

Qy 1 MISVNDK-----TGLTISVDNAIWKVI---DFQHKVPGKGSFVRSK--- 40
Db 736 IVSAQYEVRIINSEGRRELLDHSGEWINKESIIDKISSKEYISFPNKENKITVRSKNLP 795

Qy 41 -----LRLNRGA-----IOEKT-----FRAGEKVEPAMIEKNR--RMOYLYADGDNHV 81
Db 796 ELSTLLQIRNNSNSDIELEKVMLTECEINVISNIDTQIVEERIEAKNLTSDSINIY 855

Qy 82 FMNDSFEQTELSDDL-----KEELNLYKGEVEVQ-IQTYEGTIGVE 124
Db 856 ---KDEFKLIESIDALCDLQKQNELEDSHFISFEDTDEGFSIRFINKETGESIFVE 912

Qy 125 LPKTV-----ELTVTEFGIKG---DTATG-ATKSATVETGY---TLNVPLEVF 166
Db 913 TEKTIFSEYANHITTEISKIRGTFTDVTNGKLVKKVNLDTTHEVNTLNAAFFI 965

RESULT 14
US-09-021-323-3
; Sequence 3, Application US/09021323
; Patent No. 5929033
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry

; TITLE OF INVENTION: EXTRACELLULAR MUCOUS MATRIX
; TITLE OF INVENTION: GLYCOPROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,323
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0477 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 464 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 294502
; US-09-021-323-3

Query Match 8.2%; Score 77.5; DB 2; Length 464;
Best Local Similarity 22.3%; Pred. No. 2.5;
Matches 35; Conservative 22; Mismatches 45; Indels 55; Gaps 6;

Qy 9 TGLTISVDNAIWKVIDFOHKVPGKGSFVRSKLRLNRGTGAIQEKTFRAGEKVEPAMIE-- 66
Db 23 TGLAGDKHCVCVELLPDSSFPK-----RVGALEDETIRLSNRVDEDMQKLE 70

Qy 67 -----NRRMOYLYADGDNHVFMNDSFEQTELSDDLKEELNLYK---- 106
Db 71 EQDITLDITYSEKIINLTRLRVEYLEK-----LHPESL--VEISFEVLKREIRELEMYIS 121

Qy 107 -----EGMEVQIQTYEGTIGVLPKTVLTVTETE 137
Db 122 AMRVKPNNGSVQVETLYNEV-----KNMSKTVGQLE 152

RESULT 15
US-08-776-265-5
; Sequence 5, Application US/08776265
; Patent No. 6001631
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
; APPLICANT: CAMERON, Beatrice
; APPLICANT: CROUZET, Joel
; APPLICANT: FAMECHON, Alain
; APPLICANT: FERRERO, Lucia
; TITLE OF INVENTION: No. 6001631el Topoisomerase IV, Corresponding
; TITLE OF INVENTION: Nucleotide Sequences and Uses thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

```

; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/776,265
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 663 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-776-265-5

```

```

Query Match      8.0%; Score 76; DB 3; Length 663;
Best Local Similarity 26.2%; Pred. No. 6.1;
Matches 33; Conservative 22; Mismatches 29; Indels 42; Gaps 8;

Qy  8 KTG--LTISVDNAIKV---IDFQHVKG-KGSAPVRSKLRNLRGTGAIQKTFRAGEKVE 61
    ||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 163 KTGKTVTFKPDDTIFKASTSFNFDVLSERLQESAF---LKNLK---ITLNDLRSGK-- 214

Qy  62 PAMIEENRMQVLYADGDNHVPMDNESFEOTELSSDYLKEELNLYLKEGMEV--QIQTYEGE 119
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 215 -----ROEHYHBEG-----IKFVSYNMGKVELHLDVATFSGE 248

Qy 120 TIGVEL 125
    | : |
Db 249 ANGIEV 254

```

Search completed: January 12, 2001, 14:58:57
Job time: 61 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:58:40 ; Search time 16.19 Seconds
(without alignments)
369.018 Million cell updates/sec

Title: US-09-322-732-1

Perfect score: 946

Sequence: 1 MISVNDFKTGLTISVDNAIW.....VNEGDLVLIINTGDSYISRG 185

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	702	74.2	185	1 EFP_BACSU	P49778 bacillus su
2	541	57.2	185	1 EFP_SYNP7	Q54760 synechococ
3	535	56.6	187	1 EFP_SYNP3	Q55119 synechocyst
4	505	53.4	185	1 EFP_ANASP	Q44247 anabaena sp
5	480	50.7	187	1 EFP_MYCTU	P95019 mycobacteri
6	460	48.6	187	1 EFP_BRELA	Q45288 brevibacter
7	435	46.0	192	1 EFP_AQUAE	O67376 aquifex aeo
8	421	44.5	185	1 EFP_THEMEA	Q9x284 thermotoga
9	409	43.2	187	1 EFP_HELPY	P36004 helicobacte
10	407	43.0	185	1 EFP_BACFR	P70889 bacteroides
11	406	42.9	187	1 EFP_HELPJ	Q9zm95 helicobacte
12	383.5	40.5	190	1 EFP2_CHLTR	O84757 chlamydia t
13	363.5	38.4	190	1 EFP2_CHLPN	Q9x711 chlamydia p
14	363	38.4	187	1 EFP_TREPA	O83537 treponema p
15	362	38.3	187	1 EFP_ECOLI	P33398 escherichia
16	337	35.6	190	1 EFP_MYCGE	P47272 mycoplasma
17	331.5	35.0	188	1 EFP_RICPR	Q9zdt7 rickettsia
18	327	34.6	190	1 EFP_MYCPN	P75085 mycoplasma
19	311	32.9	187	1 EFP_HAEIN	P43771 haemophilus
20	284	30.0	275	1 YEIP_ECOLI	P33028 escherichia
21	280.5	29.7	187	1 EFP_BUCMP	O51834 buchnera ap
22	275	29.1	185	1 EFP1_CHLTP	O84124 chlamydia t
23	270	28.5	185	1 EFP1_CHLPN	Q9z900 chlamydia p
24	269.5	28.5	192	1 EFP_BORBU	O51232 borrelia bu
25	114.5	12.1	135	1 IF5A_SULAC	P28461 sulfolobus
26	104	11.0	132	1 IF5A_METJA	Q38625 methanococc
27	95.5	10.1	138	1 IF5A_PYRHO	O50089 pyrococcus
28	95	10.0	136	1 IF5A_PYRAE	P56635 pyrobaculum
29	94.5	10.0	386	1 DEGS_BACBR	P54663 bacillus br
30	91	9.6	280	1 DAPF_ARCFU	O29511 archaeoglob
31	91	9.6	5255	1 BKCA_BACLI	O68006 b bacitraci
32	90.5	9.6	203	1 CTC_BACSU	P14194 bacillus su
33	88	9.3	675	1 CIBA_PABPP	P57091 paenibacill

34 87 9.2 1951 1 CIN3_RAT P08104 rattus norv
35 86.5 9.1 130 1 IF5A_METH Q26955 methanobact
36 86.5 9.1 639 1 GYRB_HALSQ P21558 haloferax s
37 86.5 9.1 1174 1 C1EB_BACTA Q03745 bacillus th
38 84.5 8.9 268 1 ENGA_SALTY Q9xc18 salmonella
39 83.5 8.8 364 1 YEJB_ECOLI P33914 escherichia
40 83.5 8.8 604 1 SP20_YEAST P50875 saccharomyc
41 83.5 8.8 2179 1 POLG_HRV14 P03303 human rhino
42 83 8.8 567 1 VE06_VARV P33819 variola vir
43 82 8.7 352 1 SISI_YEAST P25294 saccharomyc
44 82 8.7 567 1 VE06_VACCC P21047 vaccinia vi
45 82 8.7 567 1 VE06_VACCV P21607 vaccinia vi

ALIGNMENTS

RESULT 1

ID EFP_BACSU STANDARD; PRT; 185 AA.
AC P49778;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Kobayashi Y., Mizuno M., Masuda S., Takemaru K., Hosono S.,
RA Sato T., Takeuchi M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 139-185 FROM N.A.
RC STRAIN=168 / JH642;
RA Guerout-Fleury A.M., Gonzy-Treboul G., Stragier P.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INVOLVED IN PEPTIDE-BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on lcs
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D84432; BAA12558.1; -;
CC EMBL; U35252; AAA76718.1; -;
CC EMBL; Z99116; CAB14376.1; -;
CC SUBTILIST; BG11460; EFP.
CC INTERPRO; IPR001059; -;
CC PFAM; PF01132; EFP; 1.
CC PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20454 MW; 60C3B9BF90D2DAB8 CRC64;

Query Match 74.28; Score 702; DB 1; Length 185;
Best Local Similarity 71.7%; Pred. No. 1.2e-48;
Matches 132; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

Qy	1	MISVNDFKTGLTISVDNAIKWVIDFQHVKPGKGSFAVRSKLRLNRGTAIQEKTFRAGEKV	60
Dd	1	MISVNDFTGLTTIDVDGGIWRVVDFQHVKPGKAFAVRSKLRLNRGTAIQEKTFRAGEKV	60
Qy	61	EPAMIENRMOYLYADGDNHIVFMNDSFEOTELSSDYLKKEELNYLKEGMEVQIQTYEGET	120
Dd	61	AKAQETKTMOYLYANGDQHVFMDTSYEQLEUSATQIBEEELYLNENSVHLMMYTODET	120
Qy	121	I GVLPRTVELTVETEPGIKGDTATGATKSATVECTYLNVPLFVNEGDVLIINTGDGS	180
Dd	121	LGIELPNTVELKVVETEPGIKGDTASGTRPAKETETCLVVNVFFVFNEDGTLVNVNTSDGS	180
Qy	181	YISR 184	
Dd	181	YYSR 184	

RESULT 2

```

EFF_SYPN7 STANDARD; PRT; 185 AA.
ID _EFN_SYPN7 AC Q54760;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
OS EFP.
GN Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
NC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
RC [1]
RP SEQUENCE FROM N.A.
RA Phung L.T., Haselkorn R.;
RT "Genes encoding biotin carboxyl carrier protein and elongation factor P from cyanobacterium Synecococcus sp. PCC 7942.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC -1- FUNCTION: INVOLVED IN PEPTIDE-BOND SYNTHESIS. STIMULATE EFFICIENT TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE (BY SIMILARITY).
CC CC
CC -1- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/annotation/send_email_to_license@isb-sib.ch).

EMBL; U59235; AAB82025.1; --
INTERPRO; IPR001059; --
PFAM; PF01132; EFP; 1.
PROSITE; PS01275; EFP; 1.
Protein biosynthesis; Elongation factor.
Sequence 185 AA; 20368 MW; BB6880A3D96CF06C CRC64;

Query Match	57.2%	Score	541;	DB	1;	Length	185;
Best Local Similarity	54.9%;	Pred.	No. 5.5e-36;				
Matches	101;	Conservative	36;	Mismatches	47;	Indels	0;
Gaps	0;						
QY	1	MISVNDPFTGLTISVDNAIKWIDFQHVKPGKGSFVRSKLRNLRNLTGAIQKETF	RAGEKV	60			
Db	1	MISVNDPFTGTTTIDGAVRVVVEFLHVKPGKGSFVRTKLKNAKTGNVVEKTF	RAGETV	60			
QY	61	EPAMIENRMOYLADGDNHVFMDNESFEOTELSSDYLKEELNLYLKEGMEVQIOTY	GET	120			
Db	61	PQAVLEKSTLQYYKKGDDPFVMDMETYEGRGLTAAIDGRVKYLKEGMEANVTW	NGQV	120			
QY	121	IGVELPKTVELTWTETEPGKIGDGTATGATKATSAIVETGYTILNVPLFVNEGDV	LIINTGDGS	180			

```

Db      121 LEVELPNSVLEVIETDPGVKGDTATGCTPAKVETGAQVMWPLFVSIGERIKIDTRNDS 180
Qy      181 YISR 184
        I : |
Db      181 YLGR 184
        I : |

RESULT 3
EFP_SYNY3 STANDARD;          PRG;    187 AA.
ID   EFP_SYNY3 Q55119;
AC   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DE   30-MAY-2000 (Rel. 39, Last annotation update)
DE   ELONGATION FACTOR P (EF-P).
EP   EFP OR SLR0434.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Suzuki T., Miyajima N.,
RA   "Sugura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- FUNCTION: INVOLVED IN PEPTIDE-BOND SYNTHESIS. STIMULATE EFFICIENT
CC   TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC   70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC   THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-tRNA, THUS INCREASING
CC   THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC   (BY SIMILARITY).
CC   -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC   This SWISS-PROT entry is copyrighted. It is produced through a collabora-
CC   between the Swiss Institute of Bioinformatics and the EMBL Outstat-
CC   the European Bioinformatics Institute. There are no restrictions on
CC   use by non-profit institutions as long as its content is in no
CC   modified and this statement is not removed. Usage by and for commer-
CC   entities requires a license agreement (See http://www.isb-sib.ch/anno-
CC   or send an email to license@isb-sib.ch)).
-----
CC   EMBL; D64001; BAA10251.1; -.
CC   INTERPRO: IPR001059; -.
DR   PFAM; PF01132; EFP; 1.
DR   PROSITE; PS01275; EFP; 1.
KW   Protein biosynthesis; Elongation factor.
SQ   SEQUENCE 187 AA; 20406 MW; B7D900A82B2B94938 CRC64;
```

[illegible]

```

RESULT 4
EFP_ANASP
ID EFP_ANASP STANDARD; PRT; 185 AA.
AC Q44247;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
RN [1]
RX MEDLINE=93352435; PubMed=8102363;
RA Gornicki P., Scappino L.A., Haselkorn R.;
RT "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
RT sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
RT protein.";
RL J. Bacteriol. 175:5268-5272(1993).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L14863; AAA74627.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20496 MW; F3F32A12CF106182 CRC64;

Query Match 53.4%; Score 505; DB 1; Length 185;
Best Local Similarity 52.2%; Pred. No. 3.7e-33;
Matches 96; Conservative 38; Mismatches 50; Indels 0; Gaps 0;

QY 1 MISVNDFTGLTISVDNAIWKVIDFQHVKPGKGSFAVRSKLNLRTGAIQKTFRAGEKV 60
DB 1 MISSNDFRPGSVILDGSWRVVIDFLHVKPGKGSFAVRTTLKNVSGKVKLEKTFRAGETV 60

QY 61 EPAMLENRMQVLYADGNHFMVDNESFEQTELSDDYLKEELNLYLKEGMEVQIQYEGET 120
DB 61 POATLEKITMQHTYKEGDFEFVMDSEYEGRLSAAQIGDRVKYKLEGVNVIWGEQV 120

QY 121 IGVLPKTVLTVTETPGIKGDTATGATKSATVTGTTLNVPFLVNGEDVLIINTGDS 180
DB 121 LVEVLANSVVLVIQTDPGVKGDTATGGTKPAIVETGATVMVPLPFIISGERIKIDTRDK 180

QY 181 YISR 184
DB 181 YLGR 184

RESULT 5
EFP_MYCTU
ID EFP_MYCTU STANDARD; PRT; 187 AA.
AC P95019;
DT 01-NOV-1997 (Rel. 35, Created)

01-NOV-1997 (Rel. 35, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
ELONGATION FACTOR P (EF-P).
EFP OR RV2534C OR MTCY159.22.
Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium.
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
[1]
SEQUENCE FROM N.A.
STRAIN-H37RV;
MEDLINE=98295987; PubMed=9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
Nature 393:537-544(1998).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z83863; CAB06174.1; -.
DR TUBERCULIST; RV2534C; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 187 AA; 20407 MW; 58BCBD1DE0424F58 CRC64;

Query Match 50.7%; Score 480; DB 1; Length 187;
Best Local Similarity 47.8%; Pred. No. 3.4e-31;
Matches 88; Conservative 38; Mismatches 58; Indels 0; Gaps 0;

QY 1 MISVNDFTGLTISVDNAIWKVIDFQHVKPGKGSFAVRSKLNLRTGAIQKTFRAGEKV 60
DB 1 MATTAADFNGNLVLDGQLWTITEFQHVKPGKGFAPFRTKLKNVLSGKVDKTFNAGKV 60

QY 61 EPAMLENRMQVLYADGNHFMVDNESFEQTELSDDYLKEELNLYLKEGMEVQIQYEGET 120
DB 61 DTATVDRTTYLYRDGSDGFVMDSQDYEQHPLPEALYGDAAARFLLEGMPQVQAFHNGVP 120

QY 121 IGVLPKTVLTVTETPGIKGDTATGATKSATVTGTTLNVPFLVNGEDVLIINTGDS 180
DB 121 LVIELPVTVELEVTHTEPLGQDRSSAGTKPATLQTAQINVPFLPINTGDKLVKVDSDGS 180

QY 181 YISR 184
DB 181 YLGR 184

RESULT 6
EFP_BRELA
ID EFP_BRELA STANDARD; PRT; 187 AA.
DT 01-NOV-1997 (Rel. 35, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
ELONGATION FACTOR P (EF-P).
EFP OR RV2534C OR MTCY159.22.
Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium.
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
[1]
SEQUENCE FROM N.A.
STRAIN-H37RV;
MEDLINE=98295987; PubMed=9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
Nature 393:537-544(1998).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z83863; CAB06174.1; -.
DR TUBERCULIST; RV2534C; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 187 AA; 20407 MW; 58BCBD1DE0424F58 CRC64;
```

```

AC Q45288;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Brevibacterium lactofermentum.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13869;
RX MEDLINE=98036051; PubMed=9370284;
RA Ramos A., Macias J.R., Gil J.A.;
RT "Cloning, sequencing and expression of the gene encoding elongation
RT factor P in the amino-acid producer Brevibacterium lactofermentum
RT (Corynebacterium glutamicum ATCC 13869).";
RL Gene 198:217-222(1997).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99289; CAA67673.1; -
DR INTERPRO: IPR001059; -
DR PFAM: PF01132; EFP; 1.
DR PROSITE: PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 187 AA; 20626 MW; 9E62823094CB0C1B CRC64;

Query Match 48.6%; Score 460; DB 1; Length 187;
Best Local Similarity 48.4%; Pred. No. 1.3e-29;
Matches 89; Conservative 38; Mismatches 57; Indels 0; Gaps 0;

Qy 1 MISVNDKFTGLTISVDNAIWKVIDFOHVKPGKSAFVRSKLRNLRTGAIQEKTFRAGEKV 60
   | : ||| || : : : ||| ||| ||| ||| ||| : || : ||| : |||
Db 1 MATTADEKNGLVKNEGLQIIEFOHVKPGKPAFVRTKLKDVVTGKTIDKTNAGVKV 60

Qy 61 EPAMTENRMOYLYADGDNHVFMDNESFEQTELSDDLKEELNYLKEGMEVQIQTYGETI 120
   | : ||| || : ||| || : ||| ||| ||| ||| : : ||| ||| : |||
Db 61 ETATVDPRCTLYLNDGTSFIVMDKTFEQVELSPDAGDAGRFLLENMRVVOVSFHGEA 120

Qy 121 IGVLPKTVELTVTETEPGIKGDTATGATSKATVETGYTLNVPFLVNEGDVLIINTGDGS 180
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 121 LFGELPVSVDLRVETDPGLQDSTGTRPATLETGAETQVPLFIETGNVLKVDTRDGS 180

Qy 181 YISR 184
   | : |||
Db 181 YLSR 184

RESULT 7
ID EFP_AQUAE STANDARD; PRT; 192 AA.
AC O67376;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).

GN EFP OR AQ_1364.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000736; AAC07331.1; -
DR INTERPRO: IPR001059; -
DR PFAM: PF01132; EFP; 1.
DR PROSITE: PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 192 AA; 21807 MW; 9B651E676507CAA0 CRC64;

Query Match 46.0%; Score 435; DB 1; Length 192;
Best Local Similarity 44.3%; Pred. No. 1.2e-27;
Matches 81; Conservative 38; Mismatches 54; Indels 0; Gaps 0;

Qy 2 ISVNDKFTGLTISVDNAIWKVIDFOHVKPGKSAFVRSKLRNLRTGAIQEKTFRAGEKVE 61
   | : ||| || : ||| || : ||| ||| ||| ||| : || : ||| : |||
Db 5 IDINRIQKDFIEHKGEYRVLDYEHVKGPGKGAQFVRSKLRNLRTGAIQEKTFRAGEKVE 64

Qy 62 PAMTENRMOYLYADGDNHVFMDNESFEQTELSDDLKEELNYLKEGMEVQIQTYGETI 121
   | : ||| || : ||| || : ||| ||| ||| ||| : : ||| ||| : |||
Db 65 LADPEQVYATYSYNDGENYVFMTQTYDMLAVPKETEEAEAKFLKEGMEVIVFLYKGOPI 124

Qy 122 GVLPKTVELTVTETEPGIKGDTATGATSKATVETGYTLNVPFLVNEGDVLIINTGDGSY 181
   | : ||| || : ||| || : ||| ||| ||| ||| : ||| ||| : |||
Db 125 GIELPKHVELOVWETEPAFKGDQAGTKPAKLETGAVIQVPPFVKEGDVKVDTRTGSY 184

Qy 182 ISR 184
   | : |||
Db 185 VER 187

RESULT 8
ID EFP_THEME STANDARD; PRT; 185 AA.
AC Q9X284;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP OR TM1763.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=MSB8 / DSM 3109;
RA MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001815; AAD36827.1; -
DR TIGR; TM1763; -
DR PROSITE; PS01275; EPP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20875 MW; A18A98DF3597A68E CRC64;

Query Match 44.5%; Score 421; DB 1; Length 185;
Best Local Similarity 44.6%; Pred. No. 1.5e-26;
Matches 82; Conservative 35; Mismatches 67; Indels 0; Gaps 0;

Qy 1 MISVNDFTGLTISVDNAIWKVIDFOHVKPGKSAFVRSKLRNLTGAIQKTRFRAGEKV 60
Db 1 MIEVGDLLKGGFIYDGEIYRVLKASKHFMGRGSLIKLKNVKTGFRVFNFPSEGEKV 60
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 61 EPAMTENRRMQLYADGDNHVFMDSNEFQTELSDDLKEELNLYLKEGMEVQIQTYEGET 120
Db 61 QEAELSFRAQYLYRDGDHYFMTLDDYEQVALSEEEIGDAKYILVENKVDLYFHEGTP 120
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 121 IGVELPKTVELTVTEPGIKGDTATGATKSAVTGTGLNVLPLFVNEGDLVLLINTGDS 180
Db 121 IGIELPTVELTVTEPSEFRKGTVSGGKPAVLETGLKITVPIEYVIEGDKIKVDTRTGE 180
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 181 YISR 184
Db 181 YVGR 184
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
EPP_HELPY
ID EPP_HELPY STANDARD; PRT; 187 AA.
AC P56004;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EPP OR HP0177.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;

```

```

RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Cocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000538; AAD07247.1; -
DR TIGR; HP0177; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EPP; 1.
DR PROSITE; PS01275; EPP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 187 AA; 20788 MW; 1B433EDF0FAB32B CRC64;

Query Match 43.2%; Score 409; DB 1; Length 187;
Best Local Similarity 40.4%; Pred. No. 1.3e-25;
Matches 74; Conservative 45; Mismatches 64; Indels 0; Gaps 0;

Qy 2 ISVNDFTGLTISVDNAIWKVIDFOHVKPGKSAFVRSKLRNLTGAIQKTRFRAGEKV 61
Db 3 IGMSELKKKLTGELGGVPYRIVEYOHVVKPGKGAAPVRAKIKSFLDGGKVIKTFHAGDKCE 62
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 62 PAMTENRRMQLYADGDNHVFMDSNEFQTELSDDLKEELNLYLKEGMEVQIQTYEGETI 121
Db 63 EPNLVEKTMQYLYHDGDTYQFMDIESYEQIALNDSQVGEASKWMLDGMQVQLLHNDKAI 122
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 122 GVLEPKTVELTVTEPGIKGDTATGATKSAVTGTGLNVLPLFVNEGDLVLLINTGDSY 181
Db 123 SVDVPPVVALKIVETAPNFKGDTSSASKKPALETGAVVQVFPFHVLEGEIILKVNTEY 182
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 182 ISR 184
Db 183 LEK 185
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
EPP_BACFR
ID EPP_BACFR STANDARD; PRT; 185 AA.
AC P70889;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EPP.
OS Bacteroides fragilis.
OC Bacteria; CFB group; Bacteroidaceae; Bacteroides.
RN [1]
RP SEQUENCE FROM N.A.

```

```
RC STRAIN=BFI;
RX MEDLINE=983111074; PubMed=9648740;
RA Abcartt V.R., Mbewe M., Woods D.R.;
RT "Cloning of an EF-P homologue from Bacteroides fragilis that
RT increases B. fragilis glutamine synthetase activity in Escherichia
RT coli.";
RL Mol. Gen. Genet. 258:363-372(1998).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U75509; AAC26328.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20812 MW; 2457150F941A240E CRC64;
```

```
Query Match 43.08; Score 407; DB 1; Length 185;
Best Local Similarity 44.18; Pred. No. 1.8e-25;
Matches 82; Conservative 44; Mismatches 34; Indels 6; Gaps 4;
```

```
Qy 1 MISVNDFTGLTISVDNAIKWVIDFOHVKPGKGSFAVRSKLNRLRTGAIQKTFRAGEKV 60
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1 MINAQDIKNGTCIRMDG---KLIEFLHVKPGKGNTEMTKLDKDVGVGVLRERFNIKEKL 57

Qy 61 EPAMIENRMOYLYADGNHVMFNDESFEQTELSDDLKEELNLYLKEGMEVQIOTVEGETI 119
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 58 EDVRRPYPYLYKGEDEYIFMNOETFDQHPAHD-LINGVDPLLEGAVVEVVSADSTE 116

Qy 120 TI-GVELPKTVELTVEPEIGKGTATGATKTSATVETGTLNVPFLFVNEGDLIINTGD 178
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 117 TVLYADMPKIVQMVTYTEPEGLKGTATNTLKPATVSGATVRVPLFISEGETEIDTRD 176

Qy 179 GSYISR 184
| : : : |
Db 177 GSYVGR 182
```

```
RESULT 11
EFP_HELPJ
ID EFP_HELPJ STANDARD; PRT; 187 AA.
AC Q92MQ5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP OR HP0177.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Gullid B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
```

```
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001454; AAD05735.1; -.
DR HSSP; PS6635; IBKB.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 187 AA; 20776 MW; 21733EDFB0FAB331 CRC64;
```

```
Query Match 42.98; Score 406; DB 1; Length 187;
Best Local Similarity 40.48; Pred. No. 2.2e-25;
Matches 74; Conservative 44; Mismatches 65; Indels 0; Gaps 0;
```

```
Qy 2 ISVNDFTGLTISVDNAIKWVIDFOHVKPGKGSFAVRSKLNRLRTGAIQKTFRAGEKVE 61
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 3 IGMSELKKGKUEIGGVYRIVEYQHVKPGKGAFAVRAKIKSFDLDGKVIKTFHAGDKCE 62

Qy 62 PAMIENRMOYLYADGNHVMFNDESFEQTELSDDLKEELNLYLKEGMEVQIOTVEGETI 121
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 63 EPNLVERMTQYLYHDGTYQFMDSYEQIALNDSDGVCEASKWMLDGMQVQLLHNDKAI 122

Qy 122 GVELPKTVELTVEPEIGKGTATGATKTSATVETGTLNVPFLFVNEGDLIINTGDSY 181
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 123 SDVYPQVVALKIVETAPNFKGDTSSASKPATLETGAVVQVVPFHVLEGETIKVNTETEY 182

Qy 182 ISR 184
| : : : |
Db 183 LEK 185
```

```
RESULT 12
EFP2_CHLTR
ID EFP2_CHLTR STANDARD; PRT; 190 AA.
AC O84757;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P 2 (EF-P 2).
GN EFP2 OR CT752.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=99008089; PubMed=9784136;
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
```



```
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -|- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -|- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001347; AAC68347.1; -.
CC DR PFAM; PF01132; EFP; 1.
CC DR PROSITE; PS01275; EFP; 1.
CC KW Protein biosynthesis; Elongation factor.
CC SQ SEQUENCE 190 AA; 21520 MW; 75FD7EBEFC20D7 CRC64;
CC -----
Query Match 40.5%; Score 383.5; DB 1; Length 190;
Best Local Similarity 41.3%; Pred. No. 1.3e-23;
Matches 76; Conservative 39; Mismatches 58; Indels 1; Gaps 1;
Qy 2 ISVNDFTGLTISVDNAIKWKVIDFOHVKPGKSAFVRSKLNRLRTGAIOETFRAGEKVE 61
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 4 VSTSEFRVGLRVKIDQPVILQNDVFVKPGKQAFNRKIKVKNFLTGRVIERFKSGESIE 63
Qy 62 PAMIEENRMVLYADGDNHVFMDNESFEQTELSDDLKEELNLYLKEGMEVQIQYEGETI 121
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 64 TADVREQMRLLYTDQEGATEMDDETPEQELIFWFKLENVRQWLEDTIYTLVLYNGDVI 123
Qy 122 GVLPKTVELVITETPEGIKGTATG-ATKSATVETGYTLNVPFVNEGDVLIINTGDGS 180
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 124 VSEPPFIMELTIAETAPGVGDTSAGRVLKPAVTNTGAKIMVPFIEEGEVKVDTRTGS 183
Qy 181 YISR 184
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 184 YESR 187
Qy 122 GVLPKTVELVITETPEGIKGTATG-ATKSATVETGYTLNVPFVNEGDVLIINTGDGS 180
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 124 VSEPPFIMELTIAETAPGVGDTSAGRVLKPAVTNTGAKIMVPFIEEGEVKVDTRTGS 183
Qy 181 YISR 184
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 184 YESR 187
RESULT 13
EFP2_CHLPN
ID EFP2_CHLPN STANDARD; PRT; 190 AA.
AC Q92711;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
GN EFP2 OR CPN0895
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CNL029;
RX MEDLINE=95206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Greenwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389(1999).
CC -|- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -|- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -|- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001670; AAD19033.1; -.
CC DR PROSITE; PS01275; EFP; 1.
CC KW Protein biosynthesis; Elongation factor.
CC SQ SEQUENCE 190 AA; 21477 MW; 1A245E50BDAA340C CRC64;
CC -----
Query Match 38.4%; Score 363.5; DB 1; Length 190;
Best Local Similarity 39.7%; Pred. No. 5e-22;
Matches 73; Conservative 41; Mismatches 69; Indels 1; Gaps 1;
Qy 2 ISVNDFTGLTISVDNAIKWKVIDFOHVKPGKSAFVRSKLNRLRTGAIOETFRAGEKVE 61
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 4 VSTSEFRVGLRIEIDGQPYLILQNDVFVKPGKQAFNRKIKVKNFLTGRVIERFKSGESVE 63
Qy 62 PAMIEENRMVLYADGDNHVFMDNESFEQTELSDDLKEELNLYLKEGMEVQIQYEGETI 121
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 64 TADIVERSMRLLYTDQEGATEMDDETPEQEVFWEKLENIRQWLEDTIYTLVLYNGDV 123
Qy 122 GVLPKTVELVITETPEGIKGTATG-ATKSATVETGYTLNVPFVNEGDVLIINTGDGS 180
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 124 AVEPPFIMELTIAETAPGVGDTSAGRVLKPAVTNTGAKIMVPFIDEGELVKVDTRTGS 183
Qy 181 YISR 184
:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 184 YESR 187
RESULT 14
EFP_TREPA
ID EFP_TREPA STANDARD; PRT; 187 AA.
AC O83537;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP OR TP0525.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=9832770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -|- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -|- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -|- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:58:00 ; Search time 39.67 Seconds
(without alignments)
546.596 Million cell updates/sec

Title: US-09-322-732-1
Perfect score: 946
Sequence: 1 MISVNDFKTGLTISVDNAIW.....VNEGDVLIINTGSGSYISRG 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues 374700

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organellae.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match %	Length	ID		
1	730	77.2	185	2 Q9K951	Q9K951 bacillus ha	
2	487.5	51.5	188	2 Q9KX09	Q9KX09 streptomyce	
3	443	46.8	185	2 Q9RV32	Q9RV32 deinococcus	
4	439	46.4	184	10 Q9M917	Q9M917 arabidopsis	
5	434	45.9	189	2 Q9PHW3	Q9PHW3 campylobact	
6	379.5	40.1	190	2 Q9PLH1	Q9PLH1 chlamydia m	
7	372	39.3	193	2 Q9PAM3	Q9PAM3 xylella fas	
8	371	39.2	188	2 Q9KNS1	Q9KNS1 vibrio chol	
9	363.5	38.4	190	2 Q9J0B4	Q9J0B4 chlamydia p	
10	343.5	36.3	188	2 Q9PQJ3	Q9PQJ3 ureaplasma	
11	332	35.1	136	2 Q9S3F5	Q9S3F5 helicobacte	
12	330	34.9	136	2 Q9S3F9	Q9S3F9 helicobacte	
13	330	34.9	136	2 Q9R3F4	Q9R3F4 helicobacte	
14	330	34.9	136	2 Q9R2M0	Q9R2M0 helicobacte	
15	328	34.7	187	2 Q9RNJ9	Q9RNJ9 zymomonas m	
16	326	34.5	136	2 Q9S3F8	Q9S3F8 helicobacte	
17	326	34.5	136	2 Q9S3F7	Q9S3F7 helicobacte	
18	326	34.5	136	2 Q9S3F6	Q9S3F6 helicobacte	
19	324	34.2	161	2 Q69281	Q69281 citrobacter	

RESULTS		ALIGNMENTS	
RESULT	ID	PRELIMINARY;	PRG;
Q9K951	Q9K951	PRELIMINARY;	PRG;
AC	Q9K951	PRELIMINARY;	PRG;
DT	Q9K951	PRELIMINARY;	PRG;
DT	Q9K951	PRELIMINARY;	PRG;
DT	Q9K951	PRELIMINARY;	PRG;
DE	Q9K951	PRELIMINARY;	PRG;
GN	Q9K951	PRELIMINARY;	PRG;
OS	Q9K951	PRELIMINARY;	PRG;
OC	Q9K951	PRELIMINARY;	PRG;
OC	Q9K951	PRELIMINARY;	PRG;
OX	Q9K951	PRELIMINARY;	PRG;
RN	Q9K951	PRELIMINARY;	PRG;
RP	Q9K951	PRELIMINARY;	PRG;
RC	Q9K951	PRELIMINARY;	PRG;
RA	Q9K951	PRELIMINARY;	PRG;
RL	Q9K951	PRELIMINARY;	PRG;
KW	Q9K951	PRELIMINARY;	PRG;
SQ	Q9K951	PRELIMINARY;	PRG;

Q9S3G0	helicobacte
Q9PKR6	chlamydia m
Q9KSP7	vibrio chol
Q9JG76	chlamydia p
Q9JUJ2	neisseria m
Q9JZQ8	neisseria m
Q9PBE1	xylella fas
P82459	streptococc
Q9STQ4	arabidopsis
Q9VOM2	pyrococcus
Q9W148	drosophila
Q9K6U6	bacillus ha
Q9XDY4	unidentifile
Q9VLZ0	drosophila
Q9VLY9	drosophila
Q9NFV7	drosophila
Q9K9M0	bacillus ha
Q9368	brassica ol
Q41213	brassica na
Q9VPW3	drosophila
Q46034	clostridium
P71472	lactobacill
Q98682	simian cyto
Q9Kf85	bacillus ha
Q52644	ruminococc
O64756	arabidopsis

Query Match	77.2%	Score	730	DB	2	Length	185
Best Local Similarity	75.0%	Pred. No.	2e-50				
Matches	138	Conservative	21	Mismatches	25	Indels	0
Gaps	0						

Qy	1	MISVNDFKTGLTISVDNAIWVIDFOHVKPKGSAFVRSKLRNLTGAIQKTRFAGEKV	60
Db	1	MISVNDFKTGLTISVDNAIWVIDFOHVKPKGSAFVRSKLRNLTGAIQKTRFAGEKV	60
Qy	61	EPAMIEHNRMOYLYADGNHVFMDNESFEQTELSDDLKKEELNLYKCEMVEVOIFQYGET	120
Db	61	SKAIEHNRMOYLYASGDVHFMNDNETFEQLELSTAQIEHLEKFEKEMVHVHVSQGET	120
Qy	121	IGVELPKTIVELTETEPGKNGDTSATVETGTYTLNVPLFVNEGDVLIINTGDS	180
Db	121	LGVEVPNTVELTETEPGKNGDTSATVETGTYTLNVPLFVNEGDVLIINTGDS	180
Qy	181	YISR	184
Db	181	YVSR	184

```
RESULT 2
ID Q9KXQ9 PRELIMINARY; PRT; 188 AA.
AC Q9KXQ9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ELONGATION FACTOR P.
GN EFP.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cardeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RL "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL357523; CAB93371.1; -.
SQ SEQUENCE 188 AA; 20649 MW; 3058EFEB64B688C1C CRC64;

Query Match 51.5%; Score 487.5; DB 2; Length 188;
Best Local Similarity 48.1%; Pred. No. 2.9e-31;
Matches 89; Conservative 38; Mismatches 57; Indels 1; Gaps 1;

QY 1 MISVNDFKTGLTISVDNA-IWKVIDFOHVKPGKGSFVRSKLRNLRTGAIQKTFRAGEK 59
Db 1 MASNDLKNGLVLEGGQLSVVVFQHVKPGKGFPAFVRTKLNVLGSKVVDKTFNAGVK 60

QY 60 VEPAMENRRMOYLYADGDNHVFMNDSFEQTELSDDLKKEELNLYLKEGMEVOIQTYEGE 119
Db 61 VETATVDKRDQFSDYMDGDYFVFMDMETFDQVHLGKNIIVSAAKFMKENTEEVAMYGDKA 120

QY 120 TIGVELPKTVELTETEPGIGKDTATGATKSAFVETGYTLNVPLFVNEGDLVLIINTGDG 179
Db 121 VLFVELPAAVELTQETEPGQVGRDSTGGTKPATLETGHQINVPFLFITTEKIKVDTRTS 180

QY 180 SYISR 184
Db 181 DYLR 185

RESULT 3
Q9RY32 PRELIMINARY; PRT; 185 AA.
AC Q9RY32;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P.
GN DR0119.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;

RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Mofat K.S., Qin H., Jiang L., Pamohile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daily M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001874; AAF09709.1; -.
DR TIGR; DR0119; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Elongation factor.
SQ SEQUENCE 185 AA; 20475 MW; 893FD0850AC102D3 CRC64;

Query Match 46.8%; Score 443; DB 2; Length 185;
Best Local Similarity 40.8%; Pred. No. 9.3e-28;
Matches 75; Conservative 49; Mismatches 60; Indels 0; Gaps 0;

QY 1 MISVNDFKTGLTISVDNAIWKVIDFOHVKPGKGSFVRSKLRNLRTGAIQKTFRAGEKV 60
Db 1 MISVTELRNCTKVQMDGGLWECLDYSHLKMGRGAKVVTFRNMESGSIIVDRTFNSTEKL 60

QY 61 EPAMENRRMOYLYADGDNHVFMNDSFEQTELSDDLKKEELNLYLKEGMEVOIQTYEGE 120
Db 61 QDTYVEGKMOYLYPDGDYFVFMDMETFDQVHLGKNIIVSAAKFMKENTEEVAMYGDKA 120

QY 121 IGVLEPKTVELTETEPGIGKDTATGATKSAFVETGYTLNVPLFVNEGDLVLIINTGDG 180
Db 121 LSISLPNQVILKIQTDPGVGRDSTGGTKPATLETGHQINVPFLFITTEKIKVDTRTG 180

QY 181 YISR 184
Db 181 YLSR 184

RESULT 4
Q9M9I7 PRELIMINARY; PRT; 184 AA.
AC Q9M9I7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE ELONGATION FACTOR P (EF-P).
GN F17014.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F17014 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012562; AAF26065.1; -.
KW Elongation factor.
SQ SEQUENCE 184 AA; 20599 MW; CFC2A79C3485BAAD CRC64;

Query Match 46.4%; Score 439; DB 10; Length 184;
Best Local Similarity 45.4%; Pred. No. 1.9e-27;
Matches 83; Conservative 39; Mismatches 61; Indels 0; Gaps 0;

QY 2 ISVNDFKTGLTISVDNAIWKVIDFOHVKPGKGSFVRSKLRNLRTGAIQKTFRAGEKVE 61
RC STRAIN=RI;
```


RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfi H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitaajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Ouaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zaco M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
RL Nature 406:151-157(2000).
DR EMBL; AE004055; AAF85271.1; -;
DR INTERPRO; IPR001059; -;
DR PFAM; PF01132; EFP; 1.
SQ SEQUENCE 193 AA; 21380 MW; 8D251D1472A37D06 CRC64;

Query Match 39.3%; Score 372; DB 2; Length 193;
Best Local Similarity 40.9%; Pred. No. 4e-22;
Matches 74; Conservative 40; Mismatches 67; Indels 0; Gaps 0;
Qy 4 VNDEKFTGLTISVDNAIWKVIDFOHVKPGKGSFAFVRSKLRNLTGAIQKTPRAGEKVEPA 63
Db 11 MNDVKNMGKILVNAEPAVITDEVKPGKGQAFTRVYRLKSRGVQEVTKMKSTDILEAA 70
Qy 64 MIENRRMOYLXADGNHVMFNDSFEQTELSDDLKEELNYLKEGMEYQIOTYEGETI 123
Db 71 DVVDTDQMYLSDGEYHFMQEFTEQVQADKNGMGAEKWLKEEQCVVTLNNGVPIGV 130
Qy 124 ELPKTVLTVTETEPGKIGDTPATGATKSATVETGYTLNVPFLVNEGDVLIINTGDGYSI 183
Db 131 QPPNPFVELKITETDPLRGDTSGGGKGPATLETGAVVRVPLFVNQDEVIKVDTRSGEYVS 190
Qy 184 R 184
Db 191 R 191
RESULT 8
Q9KNS1 PRELIMINARY; PRT; 188 AA.
AC Q9KNS1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE ELONGATION FACTOR P.
GN VC2660.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.";
RL Nature 406:477-483(2000).
DR EMBL; AE004332; AAF95801.1; -;
DR TIGR; VC2660; -;
SQ SEQUENCE 188 AA; 20576 MW; 73537EA64626B9A2 CRC64;
Query Match 39.2%; Score 371; DB 2; Length 188;
Best Local Similarity 41.0%; Pred. No. 4.6e-22;
Matches 75; Conservative 35; Mismatches 73; Indels 0; Gaps 0;
Qy 2 ISVNDKFTGLTISVDNAIWKVIDFOHVKPGKGSFAFVRSKLRNLTGAIQKTPRAGEKVE 61
Db 4 VSTNEFKGLKIMLDNEPCVILENEYKPGKGQAFNRVIRKLLTGKVLKTFKSGDTAE 63
Qy 62 PAMTENRRMOYLXADGNHVMFNDSFEQTELSDDLKEELNYLKEGMEYQIOTYEGETI 121
Db 64 VADVVDIDLDYLDNGGEYHFMNNSFEQLAADAQKAVGENAKWLVENNTCMLTLWNGNPI 123
Qy 122 GVLPKTVLTVTETEPGKIGDTPATGATKSATVETGYTLNVPFLVNEGDVLIINTGDGYSI 181
Db 124 AVTPNPFVELEVETETDPLRGDTSGGGKGPATLETGAVVRVPLFVQIGEVIKVDTRSAEY 183
Qy 182 ISR 184
Db 184 VGR 186
RESULT 9
Q9JQB4 PRELIMINARY; PRT; 190 AA.
AC Q9JQB4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ELONGATION FACTOR P (TRANSLATION ELONGATION FACTOR P).
GN EFP_2 OR CP0971.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83538;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of *Chlamydia pneumoniae* J138 from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dodson R., Gwinn M., Nelson W., Deboy R., Kolonay J.,
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AP002548; BAA9103.1; -;
DR TIGR; AF002255; AAF38751.1; -;
DR TIGR; CP0971; -;
RW Elongation factor.

```

SQ SEQUENCE 190 AA; 21477 MW; 1A245E50BDAA340C CRC64;

Query Match 38.4%; Score 363.5; DB 2; Length 190;
Best Local Similarity 39.7%; Pred. No. 1.8e-21;
Matches 73; Conservative 41; Mismatches 69; Indels 1; Gaps 1;

QY 2 ISVNDFTGLTISVDNAIKWIDFQHVKPGKGSFAVRSKLRNLTGATGAEKTFRAGEKVE 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 VSTSEFRVGLRIEIDQPYLQNDVFQKQAFNRKIKVNFRTGRIYTKSGESVE 63

QY 62 PAMIERRMQVLYADGDNHVFMDNESFQTELSDDYLKEELNLYLKEGMEVOIQTYEGTI 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
64 TADIVERSMRLLYTDQEGATMDDETFFQEVFWEKLENIQWLEDTIYTLVLYNGDV 123

QY 122 GVLPKTVLTETPEPIKGTATG-ATKSATVTGTGTVLNPLFVNGDVLIIINTGDS 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
124 AVEPIFMELSTAEATAPGVGDGTASGVKLPVNTGAKIMVPIDFIDGELVKVDTRTGS 183

QY 181 YISR 184
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
184 YISR 187

RESULT 10
Q9PQJ3 PRELIMINARY; PRT; 188 AA.
AC Q9PQJ3;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE TRANSLATION ELONGATION FACTOR P.
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplastmataceae; Ureaplasma.
OX NCBI_TaxID=2130;
RP SEQUENCE FROM N.A.
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of Ureaplasma urealyticum: Alternate views of a
RT minimal genome and sexually transmitted pathogen.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE002127; AAF30707.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
SQ SEQUENCE 188 AA; 21221 MW; 6EE4A873CF7F7E15 CRC64;

Query Match 36.3%; Score 343.5; DB 2; Length 188;
Best Local Similarity 37.0%; Pred. No. 6.8e-20;
Matches 68; Conservative 39; Mismatches 76; Indels 1; Gaps 1;

QY 1 MTSVNDFTGLTISVDNAIKWIDFQHVKPGKGSFAVRSKLRNLTGATGAEKTFRAGEKVE 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 IIOAKDLRAGHTFLYKGNLYQVNIQVENSFNKTAMREGIVKCKVKNLRTGATTEVL-TGEKL 62

QY 61 EPAMIERRMQVLYADGDNHVFMDNESFQTELSDDYLKEELNLYLKEGMEVOIQTYEGTI 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 EQAVIEKSKMTFSYDDGSGYFVMDNDTTEQISIPYSQLSWEKNFTEEGTSEVSMRYDDEL 122

QY 121 IGVELPKTVLTETPEPIKGTATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 YISR 184
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
183 YVGR 186

RESULT 11
Q9S3F5

```

```

ID Q9S3F5 PRELIMINARY; PRT; 136 AA.
AC Q9S3F5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPI;
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239624; CAB37735.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15008 MW; B317AF0A270ADA4B CRC64;

Query Match 35.1%; Score 332; DB 2; Length 136;
Best Local Similarity 45.9%; Pred. No. 3.5e-19;
Matches 61; Conservative 30; Mismatches 42; Indels 0; Gaps 0;

QY 25 FOHVKPGKGSFAVRSKLRNLTGATGAEKTFRAGEKVEPAMIERRMQVLYADGDNHVFMD 84
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 YQHVKPGKGAFAVRKIKSFLDGVKIEKTFHAGDKCEKSNLVEKTMQVLYHDGDTYQFMD 60

QY 85 NESFEQTELSDDYLKEELNLYLKEGMEVOIQTYEGTICGVLPKTVLTETPEGIKGT 144
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 IESVEQIALNDSQVGEASKWMLDGMQVQLLHNDKAKISVDVPQVVALKIVETAPNFKGDT 120

QY 145 ATGATKSATVETG 157
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 SSASKKPATLETG 133

RESULT 12
Q9S3F9 PRELIMINARY; PRT; 136 AA.
AC Q9S3F9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=#12;
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239608; CAB37719.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15049 MW; FCCA2867F2878289 CRC64;

```

```

Query Match          34.9%; Score 330; DB 2; Length 136;
Best Local Similarity 45.9%; Pred. No. 5.1e-19;
Matches 61; Conservative 29; Mismatches 43; Indels 0; Gaps 0;

Qy 25 FOHVKPGKGSFAVRSKLRNLRTGAIQKTFRAGEKVEPAMENRRMOYLYADGDNHVFMD 84
    :|||||:|||||: : : ||||| : : : ||||| : : : ||||| : : : |||||
Db 1 YQHVKPGKGAFAVRKIKSFLDKGVIEKTFHAGDKCEPNLVKTMQYLYHDGDTYQFMD 60

Qy 85 NESFEQTELSDDLKELNYLKEGMEVQIOTYEGETIGVLPKTVLTETETEPGKIGDT 144
    ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
Db 61 IESYEQIALNDSQVGEASKWMLDGMQVLLHNDKAIISVDVPQVVALKIVETAPNFKGDT 120

Qy 145 ATGATKSATVETG 157
    : : : |||||
Db 121 SSASKKPATLETG 133

RESULT 13
Q9R3F4 PRELIMINARY; PRT; 136 AA.
AC Q9R3F4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RC SEQUENCE FROM N.A.
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
  Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
  "Recombination and clonal groupings within Helicobacter pylori from
  different geographic regions.";
  Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239623; CAB37734.1; -
DR EMBL; AJ239613; CAB37724.1; -
DR EMBL; AJ239614; CAB37725.1; -
DR EMBL; AJ239615; CAB37726.1; -
DR EMBL; AJ239616; CAB37727.1; -
DR EMBL; AJ239617; CAB37728.1; -
DR EMBL; AJ239618; CAB37729.1; -
DR EMBL; AJ239619; CAB37730.1; -
DR EMBL; AJ239620; CAB37731.1; -
DR EMBL; AJ239621; CAB37732.1; -
DR EMBL; AJ239622; CAB37733.1; -
DR INTERPRO; IPR001059; -
GN EFP.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
FT SEQUENCE 136 AA; 15018 MW; 7E4970D8D2896C92 CRC64;

```

```

Query Match          34.9%; Score 330; DB 2; Length 136;
Best Local Similarity 45.9%; Pred. No. 5.1e-19;
Matches 61; Conservative 29; Mismatches 43; Indels 0; Gaps 0;

Qy 25 FOHVKPGKGSFAVRSKLRNLRTGAIQKTFRAGEKVEPAMENRRMOYLYADGDNHVFMD 84
    :|||||:|||||: : : ||||| : : : ||||| : : : ||||| : : : |||||
Db 1 YQHVKPGKGAFAVRKIKSFLDKGVIEKTFHAGDKCEPNLVKTMQYLYHDGDTYQFMD 60

Qy 85 NESFEQTELSDDLKELNYLKEGMEVQIOTYEGETIGVLPKTVLTETETEPGKIGDT 144
    ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
Db 61 IESYEQIALNDSQVGEASKWMLDGMQVLLHNDKAIISVDVPQVVALKIVETAPNFKGDT 120

Qy 145 ATGATKSATVETG 157
    : : : |||||
Db 121 SSASKKPATLETG 133

RESULT 14
Q9R2M0 PRELIMINARY; PRT; 136 AA.
ID Q9R2M0
AC Q9R2M0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

```

```

DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RC SEQUENCE FROM N.A.
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
  Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
  "Recombination and clonal groupings within Helicobacter pylori from
  different geographic regions.";
  Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239623; CAB37734.1; -
DR EMBL; AJ239613; CAB37724.1; -
DR EMBL; AJ239614; CAB37725.1; -
DR EMBL; AJ239615; CAB37726.1; -
DR EMBL; AJ239616; CAB37727.1; -
DR EMBL; AJ239617; CAB37728.1; -
DR EMBL; AJ239618; CAB37729.1; -
DR EMBL; AJ239619; CAB37730.1; -
DR EMBL; AJ239620; CAB37731.1; -
DR EMBL; AJ239621; CAB37732.1; -
DR EMBL; AJ239622; CAB37733.1; -
DR INTERPRO; IPR001059; -
GN EFP.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
FT SEQUENCE 136 AA; 15019 MW; FCC35872F287829C CRC64;

Query Match          34.9%; Score 330; DB 2; Length 136;
Best Local Similarity 45.9%; Pred. No. 5.1e-19;
Matches 61; Conservative 29; Mismatches 43; Indels 0; Gaps 0;

Qy 25 FOHVKPGKGSFAVRSKLRNLRTGAIQKTFRAGEKVEPAMENRRMOYLYADGDNHVFMD 84
    :|||||:|||||: : : ||||| : : : ||||| : : : ||||| : : : |||||
Db 1 YQHVKPGKGAFAVRKIKSFLDKGVIEKTFHAGDKCEPNLVKTMQYLYHDGDTYQFMD 60

Qy 85 NESFEQTELSDDLKELNYLKEGMEVQIOTYEGETIGVLPKTVLTETETEPGKIGDT 144
    ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
Db 61 IESYEQIALNDSQVGEASKWMLDGMQVLLHNDKAIISVDVPQVVALKIVETAPNFKGDT 120

Qy 145 ATGATKSATVETG 157
    : : : |||||
Db 121 SSASKKPATLETG 133

RESULT 15
Q9RNJ9 PRELIMINARY; PRT; 187 AA.
ID Q9RNJ9
AC Q9RNJ9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P EFP.
GN EFP.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;
OC Zymomonas.
OX NCBI_TaxID=542;
RN [1]
RC SEQUENCE FROM N.A.
RA Lee H.J., Kang H.S.;
  "Sequence analysis of 42P4 fosmid clone of Zymomonas mobilis ZM4.";
  Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF180145; AAD56928.1; -
DR INTERPRO; IPR001059; -
GN EFP.

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: January 12, 2001, 14:58:32 ; Search time 27.86 Seconds
(without alignments)
230.741 Million cell updates/sec

Title: US-09-322-732-2
Perfect score: 975
Sequence: 1 MATYNSNDFRAGLKIMLDGE.....QIGEVKIVDTRSGYVSRVK 188
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36.*
1: /SIDSI1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES											
Result No.	Score	%		Query Match	Length	DB	ID	Description			
		Score	Match								
1	454.5	46.6	190	20	Y37395			Protein involved i			
2	445.5	45.7	190	20	Y35550			Chlamydia pneumoni			
3	362	37.1	185	21	Y70152			Staphylococcus aur			
4	229	23.5	185	20	Y37084			Protein involved i			
5	226	23.2	185	20	Y34781			Chlamydia pneumoni			
6	78	8.0	605	20	Y16791			Truncated plant-op			
7	78	8.0	718	19	W61024			Chimeric lepidopte			
8	78	8.0	1148	16	R84731			CryIF/cryIA(b) chi			
9	78	8.0	1148	16	R84733			CryIF/cryIA(b) chi			
10	78	8.0	1148	17	R94907			CryIF/cryIA(b) chi			
11	78	8.0	1148	19	W76718			Alternative cryIF/			
12	78	8.0	1148	19	W76717			Alternative cryIF/			

13	78	8.0	1148	19	W76707	Plasmid pMYC2244 c
14	78	8.0	1148	19	W76709	Plasmid pMYC2523 c
15	78	8.0	1148	19	W76715	Consensus cryIF/cr
16	78	8.0	1148	20	Y16790	Plant-optimised cr
17	78	8.0	1150	16	R84734	CryIA(c)/CryIF/cry
18	78	8.0	1150	17	R94913	CryIA(c)/CryIF/cry
19	78	8.0	1150	19	W76706	Plasmid pMYC2239 b
20	78	8.0	1174	12	R10131	Lepidopteran-activ
21	78	8.0	1174	12	R14856	CryIF protein. Ba
22	78	8.0	1174	14	R39754	Delta endotoxin.
23	78	8.0	1174	16	R84735	CryIF toxin with l
24	78	8.0	1174	16	R84732	CryIF/436 chimeric
25	78	8.0	1174	17	R94908	CryIF/436 chimeric
26	78	8.0	1174	17	R94914	CryIF toxin encode
27	78	8.0	1174	17	R94916	Native CryIF/CryIA
28	78	8.0	1174	19	W76708	Plasmid pMYC2243 B
29	78	8.0	1174	19	W76710	Plasmid pMYC2254 c
30	78	8.0	1174	19	W76712	B. thuringiensis c
31	78	8.0	1174	20	Y16792	Wild-type B.t cryI
32	78	8.0	1175	19	W76716	Alternative cryIF/
33	78	8.0	1184	16	R84737	Native cryIF toxin
34	77	7.9	609	21	Y58924	E. coli glucosamin
35	76	7.8	609	21	Y58822	E. coli glucosamin
36	76	7.8	609	21	Y58823	E. coli glucosamin
37	76	7.8	609	21	Y58825	E. coli glucosamin
38	76	7.8	609	21	Y58826	E. coli glucosamin
39	76	7.8	609	21	Y58827	E. coli glucosamin
40	75	7.7	170	21	Y59205	Caldicellulosirupt
41	74	7.6	613	21	Y92512	Human OXRE-9. Ho
42	73.5	7.5	1098	17	R99392	Haemophilus adhesi
43	73	7.5	748	21	Y70231	Human RNA-associat
44	73	7.5	1215	21	Y55791	Superheat-resistan
45	72.5	7.4	299	18	W55450	H. pylori ORF 02ae

ALIGNMENTS

RESULT 1	
Y37395	
ID Y37395 standard; Protein; 190 AA.	
XX	
AC Y37395;	
XX	
DT 07-OCU-1999 (first entry)	
XX	
DE Protein involved in transcription, translation and/or maturation.	
XX	
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;	
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;	
KW nongonococcal urethritis; epidymitis; cervicitis; salpingitis;	
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.	
XX	
OS Chlamydia trachomatis.	
XX	
PN WO9928475-A2.	
XX	
PD 10-JUN-1999.	
XX	
PF 27-NOV-1998; 98WO-IB01939.	
XX	
PR 04-NOV-1998; 98US-0107077.	
PR 28-NOV-1997; 97FR-0015041.	
PR 17-DEC-1997; 97FR-0016034.	
XX	
PA (GEST) GENSET.	
XX	
PI Griffais R;	
XX	
DR WPI; 1999-371125/31.	
XX	
PT Genome sequence of Chlamydia trachomatis	
XX	

PS Disclosure; Page 1104; 1755pp; English.

CC Y36754-Y37949 are encoded by open reading frames (ORFs) of the genome of
CC Chlamydia trachomatis (see 201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC perihepatitis, Bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.

XX Sequence 190 AA;

Query Match 46.6%; Score 454.5; DB 20; Length 190;
Best Local Similarity 48.9%; Pred. No. 1.3e-42;

Matches 92; Conservative 36; Mismatches 59; Indels 1; Gaps 1;

QY 1 MATYISNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRRLTGTTRVEKTFKSTD 60

Db 1 mrvstsefrvgrlveidgpyvlqndfvkpgkgqafnrkknftrgrviekfksge 60

QY 61 SAEGADVDDMLTYLYNDGEFHFMMNTEFEQLSADAKAIGDNKWLDDQAEICIVTLWNG 120

Db 61 sietadvregmrllytdqegatfmdetfeqelfwcklenvrqwlledtiytlvlyng 120

QY 121 QPISVTPPNFVELEIVTDPLGKGDTA-GTGGKPKATLSTGAVKVPPLFVQIGEVKVDTR 179

Db 121 dvisveppimeltiaetapgrgdtagrvtgkpatntgklnvpifieggevkvdtr 180

QY 180 SGEVVSrv 187

Db 181 tgsyesrv 188

RESULT 2

ID Y35550 standard; Protein; 190 AA.

XX AC Y35550;

XX DT 13-SEP-1999 (first entry)

XX DE Chlamydia pneumoniae transmembrane protein sequence.

XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX KW vaccine; neutralising epitope.

XX OS Chlamydia pneumoniae.

XX PN WO9927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB01890.

XX PR 04-NOV-1998; 98US-0107078.

XX PR 21-NOV-1997; 97FR-0014673.

XX PA (GEST) GENSET.

XX PI Griffais R;

XX WPI; 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

XX Page 1297; Disclosure; 1912pp; English.

XX

CC Y34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see X91900) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.

XX Sequence 190 AA;

Query Match 45.7%; Score 445.5; DB 20; Length 190;

Best Local Similarity 46.3%; Pred. No. 1.3e-41;

Matches 87; Conservative 42; Mismatches 58; Indels 1; Gaps 1;

QY 1 MATYISNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRRLTGTTRVEKTFKSTD 60

Db 1 mrvstsefrvgrlveidgpyvlqndfvkpgkgqafnrkknftrgrviekfksge 60

QY 61 SAEGADVDDMLTYLYNDGEFHFMMNTEFEQLSADAKAIGDNKWLDDQAEICIVTLWNG 120

Db 61 svetadiversmrllytdqegatfmdetfeqevfweklenirgwlledtiytlvlyng 120

QY 121 QPISVTPPNFVELEIVTDPLGKGDTA-GTGGKPKATLSTGAVKVPPLFVQIGEVKVDTR 179

Db 121 dvvaveppimeltiaetapgrgdtagrvtgkpatntgklnvpifieggevkvdtr 180

QY 180 SGEVVSrv 187

Db 181 tgsyesrv 188

RESULT 3

ID Y70152 standard; Protein; 185 AA.

XX AC Y70152;

XX DT 06-JUN-2000 (first entry)

XX DE Staphylococcus aureus protein (HGS040).

XX KW Vaccine; antibacterial; prevention; attenuation; detection;
XX KW Staphylococcal infection; neonatal conjunctivitis; skin infection;
XX KW toxic shock syndrome; osteomyelitis.

XX OS Staphylococcus aureus.

XX Key Location/Qualifiers

FT Domain 30..32 /label= Antigenic_epitope

FT Domain 76..78 /label= Antigenic_epitope

FT Domain 76..78 /label= Antigenic_epitope

XX WO200012678-A2.

XX PN 09-MAR-2000.

XX PF 31-AUG-1999; 99WO-US19726.

XX PR 01-SEP-1998; 98US-0098964.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Bailey CC, Choi GH;

XX WPI; 2000-237864/20.

XX N-PSDB; Z51235.

XX

DR	N-PSDB; T05249.
XX	
XX	Nucleic acid encoding chimeric Bacillus thuringiensis
PT	delta-endotoxin - providing increased expression in Pseudomonas,
PT	esp. for control of lepidoptera pests.
XX	
XX	Claim 6; Page 42-45; 9lpp; English.
PS	
PS	The sequence represents the cryIF/cryIA(b) chimeric toxin encoded
XX	by plasmid pMYC2244. The transition from cryIF core toxin
CC	portion to heterologous protoxin portion occurs after the
CC	characteristic CryI toxin sequence of R84729 and before the end
CC	of peptide R84730. The chimeric toxin is expressed in Pseudomonas
CC	fluorescens better than native delta endotoxins. Host cells
CC	expressing the chimeric gene and producing chimeric toxin may be
CC	used in insecticide compositions. Where the host cells are plant
CC	cells, the gene confers insect resistance to the transformed
XX	plant.
XX	
SQ	Sequence 1148 AA;
	Query Match 8.0%; Score 78; DB 16; Length 1148;
	Best Local Similarity 23.4%; Pred.No. 9.1;
	Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps
QY	18 DGPBPVAYEASE--FVKPGKGQAFARVKLRRL---TCTRVEKTFKSDSAEGADVDDML 72
Db	: :
351 dprpfyrtlsdpvrvrggfnphylvglrvaftqgtghtrfrnsgtidsideippg- 409	: :
QY	73 TYLYNDGEFW----HFNNFTFFQLSADAKAIGDNKWLDDQAEIVTLNGQPISVTPP 128
Db	: :
410 ---dmsgapwndyshvhvtfvrwge---isgsds-----rapmfswthr--satpt 456	: :
QY	129 NFVELEIVDTDPGLKGDFACTGCKPATLSTGAHVKKVPPLFVOICEVIK 175
Db	: :
457 ntldperitqlpvkahtlqsg-----ttvrrgpgftg-gdlir 494	: :
RESULT 9	
R84733	
ID	R84733 standard; Protein; 1148 AA.
XX	
AC	R84733;
XX	
DT	27-MAY-1996 (first entry)
XX	
DE	CryIF/cryIA(b) chimeric toxin.
XX	
KW	Delta endotoxin; crystal protein; chimeric toxin; insecticide;
KW	pesticide; pseudomonas fluorescens; biological control agent;
KW	transgenic plant; insect resistance; disease resistance;
KW	crop improvement; protein engineering.
XX	
OS	Bacillus thuringiensis.
XX	
FH	key Location/Qualifiers
FT	Peptide 590..601
FT	/note= "characteristic sequence of cryI toxins"
FT	Misc-difference 619 /note= "E to A substitution"
FT	Misc-difference 624 /note= "S to I substitution"
FT	Misc-difference 629 /note= "L to I substitution"
FT	Peptide 1035..1042
FT	/note= "transition sequence of chimeric toxin"
XX	
PN	WO9530753-A1.
XX	
PD	16-NOV-1995.
XX	
PF	05-MAY-1995; 95WO-US05431.
PF	


```

XX OS Synthetic.
XX PN US5827514-A.
XX PD 27-OCT-1998.
XX PF 08-FEB-1996; 96US-0598305.
XX PR 06-DEC-1994; 94US-0349867.
XX PR 08-FEB-1996; 96US-0598305.
XX PA (MYCO ) MYCOGEN CORP.
XX PI Bradfisch GA, Schwab GE, Thompson M;
XX DR WPI; 1998-593944/50.
XX PT Composition for biological control of lepidopteran pests -
XX PT comprising cells expressing two chimeric Bacillus thuringiensis
XX PT crystal proteins
XX PS Claim 10; Column 117-124; 75pp; English.
XX CC This sequence represents a Bt endotoxin which is used in a method for
XX CC controlling lepidopteran pests. The method involves the use of cells
XX CC that express the Bacillus thuringiensis delta-endotoxins cryIF chimeric
XX CC core toxin-containing protein and CryIA(c) chimeric core toxin-containing
XX CC protein in a combination that have synergistically enhanced activity,
XX CC against e.g. corn earworm (Heliothis zea).
XX SQ Sequence 1148 AA;

Query Match 8.0%; Score 78; DB 19; Length 1148;
Best Local Similarity 23.4%; Pred. No. 9.1;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;

QY 18 DGEPYAVEASE--FVKPGKGQAFARVKLRLL---TGTRVEKTFKSTDSAGADVVDNML 72
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 351 dprpfyrtlsdpvfrvggnphylgrvafqgtgnhttrfnsgrtdsldeippg- 409

QY 73 TYLYNDGEFW---HFMNNETEQLSADAKAIGDNKWLDDQAEICIVTLWNGQPISVTPP 128
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 410 ---dsggapwndyshvlnhvtfrwpg---isgsdsw-----rapmfswthr--satpt 456

QY 129 NFVELEIVDTDPGLKGDGTAGTGKPKATLSTGAVVKVPLFVQIGEVIK 175
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 457 ntldperitqiplvkahtlqsg-----ttvvrpgpgftg-gdlir 494

RESULT 12
W76717
ID W76717 standard; Protein; 1148 AA.
AC W76717;
DT 13-JAN-1999 (first entry)
DE Alternative cryIF/cryIA(b) chimeric endotoxin protein #2.
KW Bt toxin; lepidoptera; pest; cryIF; chimeric; core toxin; cryIA(c);
KW synergism; plant; primer; endotoxin.
XX Synthetic.
XX OS US5827514-A.
XX PN 27-OCT-1998.
XX PF 08-FEB-1996; 96US-0598305.
XX PR 06-DEC-1994; 94US-0349867.
XX PR 08-FEB-1996; 96US-0598305.
XX PA (MYCO ) MYCOGEN CORP.
XX PI Bradfisch GA, Schwab GE, Thompson M;
XX DR WPI; 1998-593944/50.
XX PT Composition for biological control of lepidopteran pests -
XX PT comprising cells expressing two chimeric Bacillus thuringiensis
XX PT crystal proteins
XX PS Claim 10; Column 117-124; 75pp; English.
XX CC This sequence represents a Bt endotoxin which is used in a method for
XX CC controlling lepidopteran pests. The method involves the use of cells
XX CC that express the Bacillus thuringiensis delta-endotoxins cryIF chimeric
XX CC core toxin-containing protein and CryIA(c) chimeric core toxin-containing
XX CC protein in a combination that have synergistically enhanced activity,
XX CC against e.g. corn earworm (Heliothis zea).
XX SQ Sequence 1148 AA;

Query Match 8.0%; Score 78; DB 19; Length 1148;
Best Local Similarity 23.4%; Pred. No. 9.1;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;

QY 18 DGEPYAVEASE--FVKPGKGQAFARVKLRLL---TGTRVEKTFKSTDSAGADVVDNML 72
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 351 dprpfyrtlsdpvfrvggnphylgrvafqgtgnhttrfnsgrtdsldeippg- 409

QY 73 TYLYNDGEFW---HFMNNETEQLSADAKAIGDNKWLDDQAEICIVTLWNGQPISVTPP 128
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 410 ---dsggapwndyshvlnhvtfrwpg---isgsdsw-----rapmfswthr--satpt 456

QY 129 NFVELEIVDTDPGLKGDGTAGTGKPKATLSTGAVVKVPLFVQIGEVIK 175
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 457 ntldperitqiplvkahtlqsg-----ttvvrpgpgftg-gdlir 494

RESULT 13
W76707
ID W76707 standard; Protein; 1148 AA.
AC W76707;
DT 13-JAN-1999 (first entry)
DE Plasmid pMYC2244 cryIF/cryIA(b) chimeric toxin protein fragment.
KW Bt toxin; lepidoptera; pest; cryIF; chimeric; core toxin; cryIA(c);
KW synergism; plant; primer; endotoxin.
XX Synthetic.
XX OS US5827514-A.
XX PN 27-OCT-1998.
XX PF 08-FEB-1996; 96US-0598305.
XX PR 06-DEC-1994; 94US-0349867.
XX PR 08-FEB-1996; 96US-0598305.
XX PA (MYCO ) MYCOGEN CORP.
XX PI Bradfisch GA, Schwab GE, Thompson M;
XX DR WPI; 1998-593944/50.
XX DR N-PSDB; V62080.
XX PT Composition for biological control of lepidopteran pests -

```

```

PR 08-FEB-1996; 96US-0598305.
XX (MYCO ) MYCOGEN CORP.
XX Bradfisch GA, Schwab GE, Thompson M;
XX WPI; 1998-593944/50.
XX PT Composition for biological control of lepidopteran pests -
XX PT comprising cells expressing two chimeric Bacillus thuringiensis
XX PT crystal proteins
XX PS Claim 10; Column 111-118; 75pp; English.
XX CC This sequence represents a Bt endotoxin which is used in a method for
XX CC controlling lepidopteran pests. The method involves the use of cells
XX CC that express the Bacillus thuringiensis delta-endotoxins cryIF chimeric
XX CC core toxin-containing protein and CryIA(c) chimeric core toxin-containing
XX CC protein in a combination that have synergistically enhanced activity,
XX CC against e.g. corn earworm (Heliothis zea).
XX SQ Sequence 1148 AA;

Query Match 8.0%; Score 78; DB 19; Length 1148;
Best Local Similarity 23.4%; Pred. No. 9.1;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;

QY 18 DGEPYAVEASE--FVKPGKGQAFARVKLRLL---TGTRVEKTFKSTDSAGADVVDNML 72
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 351 dprpfyrtlsdpvfrvggnphylgrvafqgtgnhttrfnsgrtdsldeippg- 409

QY 73 TYLYNDGEFW---HFMNNETEQLSADAKAIGDNKWLDDQAEICIVTLWNGQPISVTPP 128
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 410 ---dsggapwndyshvlnhvtfrwpg---isgsdsw-----rapmfswthr--satpt 456

QY 129 NFVELEIVDTDPGLKGDGTAGTGKPKATLSTGAVVKVPLFVQIGEVIK 175
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 457 ntldperitqiplvkahtlqsg-----ttvvrpgpgftg-gdlir 494

RESULT 13
W76707
ID W76707 standard; Protein; 1148 AA.
AC W76707;
DT 13-JAN-1999 (first entry)
DE Plasmid pMYC2244 cryIF/cryIA(b) chimeric toxin protein fragment.
KW Bt toxin; lepidoptera; pest; cryIF; chimeric; core toxin; cryIA(c);
KW synergism; plant; primer; endotoxin.
XX Synthetic.
XX OS US5827514-A.
XX PN 27-OCT-1998.
XX PF 08-FEB-1996; 96US-0598305.
XX PR 06-DEC-1994; 94US-0349867.
XX PR 08-FEB-1996; 96US-0598305.
XX PA (MYCO ) MYCOGEN CORP.
XX PI Bradfisch GA, Schwab GE, Thompson M;
XX DR WPI; 1998-593944/50.
XX DR N-PSDB; V62080.
XX PT Composition for biological control of lepidopteran pests -

```



```
Oy 18 DGEYAVEASE--FVKPGKGQAFARVKLRLL---TGTRVEKTEFKSTDSAEAGDVVDMNL 72
Db 351 dprpfyrtlsdgvfvgfghphvvlgrgvafqgtgnhtrtfrnsgtldslgeippq- 409
Oy 73 TYLYNDGEFW---HFMMNETFEQLSADAKAIGDQAKWLLDQABECIVTLWNGQPISVTPP 128
Db 410 ---dnsgapwndyshlnhvtfvwpge---lsgsdsw-----rapmfswthr--satpt 456
Oy 129 NFVELEIVDTPDGLKGDPTAGTGGKPKATLSTGAVVKVPLFVOIGEVIK 175
Db 457 ntldperitqiplvkahtlqsg-----ttvvrgpgftg-gdilr 494
```

Search completed: January 12, 2001, 14:58:35
Job time: 39 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:58:57 ; Search time 21.21 Seconds
(without alignments)
159.167 Million cell updates/sec

Title: US-09-322-732-2

Perfect score: 975
Sequence: 1 MATYYSNDFRAGLKIMLDGE.....QICEVIKVDTRSGEYVSRVK 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

```
Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	78	8.0	718	2	US-08-731-079A-1	Sequence 1, Appli
2	78	8.0	1148	2	US-08-349-867-23	Sequence 23, Appl
3	78	8.0	1148	1	US-08-349-867-27	Sequence 27, Appl
4	78	8.0	1148	1	US-08-239-476-23	Sequence 23, Appl
5	78	8.0	1148	1	US-08-239-476-27	Sequence 27, Appl
6	78	8.0	1148	2	US-08-598-305A-23	Sequence 23, Appl
7	78	8.0	1148	2	US-08-598-305A-27	Sequence 27, Appl
8	78	8.0	1148	2	US-08-598-305A-35	Sequence 35, Appl
9	78	8.0	1148	2	US-08-598-305A-37	Sequence 37, Appl
10	78	8.0	1148	2	US-08-598-305A-38	Sequence 38, Appl
11	78	8.0	1148	2	US-08-639-923A-23	Sequence 23, Appl
12	78	8.0	1148	2	US-08-639-923A-27	Sequence 27, Appl
13	78	8.0	1148	2	US-08-639-923A-35	Sequence 35, Appl
14	78	8.0	1148	2	US-08-639-923A-37	Sequence 37, Appl
15	78	8.0	1148	2	US-08-639-923A-38	Sequence 38, Appl
16	78	8.0	1148	4	PCT-US95-05431-23	Sequence 23, Appl
17	78	8.0	1148	4	PCT-US95-05431-27	Sequence 27, Appl
18	78	8.0	1150	1	US-08-349-867-21	Sequence 21, Appl
19	78	8.0	1150	1	US-08-239-476-21	Sequence 21, Appl
20	78	8.0	1150	2	US-08-598-305A-21	Sequence 21, Appl
21	78	8.0	1150	2	US-08-639-923A-21	Sequence 21, Appl
22	78	8.0	1150	4	PCT-US95-05431-21	Sequence 21, Appl
23	78	8.0	1174	1	US-07-828-788A-10	Sequence 10, Appl
24	78	8.0	1174	1	US-08-349-867-25	Sequence 25, Appl
25	78	8.0	1174	1	US-08-349-867-29	Sequence 29, Appl
26	78	8.0	1174	1	US-08-349-867-32	Sequence 32, Appl
27	78	8.0	1174	1	US-08-239-476-25	Sequence 25, Appl
28	78	8.0	1174	1	US-08-239-476-29	Sequence 29, Appl

Query Match 8.0%; Score 78; DB 2; Length 718;
Best Local Similarity 23.4%; Pred. No. 1.6;

QY 18 DGEYAVEASE--FVKPGKGQAFARVKLRLL--TGTRVEKTFKSTDSAEGADVDMNL 72

ALIGNMENTS

```

1
RESULT
US-08-731-079A-1
; Sequence 1, Application US/08731079A
; Patent No. 5965428
; GENERAL INFORMATION:
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Baum, James
; TITLE OF INVENTION: CHIMERIC LEPIDOPTERAN-TOXIC CRYSTAL
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5965428rls
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731.079A
; FILING DATE: 09-OCT-1996

```

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: ECG-0008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-731-079A-1

```

Sequence 8, Appl
Sequence 25, Appl
Sequence 29, Appl
Sequence 32, Appl
Sequence 25, Appl
Sequence 29, Appl
Sequence 8, Appl
Sequence 2, Appl
Sequence 10, Appl
Sequence 25, Appl
Sequence 29, Appl
Patent No. 5183600
Sequence 36, Appl
Sequence 36, Appl
Sequence 32, Appl
Sequence 32, Appl
Sequence 32, Appl


```

; INFORMATION FOR SEQ ID NO: 23:
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1148 amino acids
;   TYPE: amino acid

```


QY 129 NEVELEIVDTDPLGKSDTAGTGKGPATLSTGAVVAVPLFVQIGEVK 175
Db 457 NTIDPERITQIPLVKAHTLQSG-----TTVVRGPGFTG-GDILR 494

RESULT 9

US-08-598-305A-37
; Sequence 37, Application US/08598305A
; Patent No. 5827514
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Mark
; APPLICANT: SCHWAB, George E.
; TITLE OF INVENTION: No. 5827514el Pesticidal Compositions
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/598,305A
; FILING DATE: 08-FEB-1996
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/349,867
; FILING DATE: 06-DEC-1994
; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA86.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-598-305A-37

Query Match 8.0%; Score 78; DB 2; Length 1148;
Best Local Similarity 23.4%; Pred. No. 3.2;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;

QY 18 DGEYAVEASE--FVKPGKGQAFARVKLRLL---TGTREKTFKSTDSAGADVVDML 72
Db 351 DPRPFYRTLSDPVFVRGGFGNPHYVLGRGVAFOQTGNTHTFRNSGTIDSLDEIPQ- 409

QY 73 TLYNDGEFW---HFMNNETFEQLSADAKAIGDNKWLDDQAEICIVTLWNGQPISVTPP 128
Db 410 ---DMSGAPWNDYSHVLNHTVVRWPG---ISGDSW-----RAPMFSWTHR--SATPT 456

QY 129 NEVELEIVDTDPLGKSDTAGTGKGPATLSTGAVVAVPLFVQIGEVK 175
Db 457 NTIDPERITQIPLVKAHTLQSG-----TTVVRGPGFTG-GDILR 494

RESULT 10

US-08-598-305A-38
; Sequence 38, Application US/08598305A
; Patent No. 5827514

; GENERAL INFORMATION:
; APPLICANT: BRADFISCH, Gregory A.
; APPLICANT: THOMPSON, Mark
; APPLICANT: SCHWAB, George E.
; TITLE OF INVENTION: No. 5827514el Pesticidal Compositions
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/598,305A
; FILING DATE: 08-FEB-1996
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/349,867
; FILING DATE: 06-DEC-1994
; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA86.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-598-305A-38

Query Match 8.0%; Score 78; DB 2; Length 1148;
Best Local Similarity 23.4%; Pred. No. 3.2;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;

QY 18 DGEYAVEASE--FVKPGKGQAFARVKLRLL---TGTREKTFKSTDSAGADVVDML 72
Db 351 DPRPFYRTLSDPVFVRGGFGNPHYVLGRGVAFOQTGNTHTFRNSGTIDSLDEIPQ- 409

QY 73 TLYNDGEFW---HFMNNETFEQLSADAKAIGDNKWLDDQAEICIVTLWNGQPISVTPP 128
Db 410 ---DMSGAPWNDYSHVLNHTVVRWPG---ISGDSW-----RAPMFSWTHR--SATPT 456

QY 129 NEVELEIVDTDPLGKSDTAGTGKGPATLSTGAVVAVPLFVQIGEVK 175
Db 457 NTIDPERITQIPLVKAHTLQSG-----TTVVRGPGFTG-GDILR 494

RESULT 11

US-08-639-923A-23
; Sequence 23, Application US/08639923A
; Patent No. 5840554
; GENERAL INFORMATION:

; APPLICANT: Thompson, Mark
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
; TITLE OF INVENTION: Pseudomonas fluorescens
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1


```

;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1148 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
US-08-639-923A-35

```

	Query Match	8.0%;	Score 78;	DB 2;	Length 1148;
	Best Local Similarity	23.4%;	Pred. No. 3.2;		
	Matches	39;	Conservative	29;	Mismatches 67; Indels 32; Gaps 9;
Qy	18	DGPYAYEASE--FVKGPKGAQFARVKLRLL---	TCTREYKTFKSTDSAGADVDMNL	72	
Db	351	DPRFFYTLSDPVFVRGGFGNPHVGLGRVAFQQTCTNHTRTFRNSGTIDSLDEIPQ-	409		
Qy	73	TYLYNDGEFN----HFNNNTFFEOLSADAKAIGDNKALLDQAEICVTLANGQOISVTPP	128		
Db	410	--DNSCAPNDYSHVLNHHVTFVRWPGC-----ISGSDSW-----RAPFSTWTR--SATPT	456		
Qy	129	NFVELEIVTDPLGLKGTACTGKGPATLSTGAVVKVPLFVQIGEVK	175		
Db	457	NTIDPERITOIPLVKAITLQSG-----TTVVRGGEGFTG-GDILR	494		

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,923A
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,476
FILING DATE: 06-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1148 amino acids

Query Match	8.0%;	Score 78;	DB 2;	Length 1148;
Best Local Similarity	23.4%;	Pred. No. 3.2;		

```

Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;
Qy 18 DGEYVAVEASE--FVKPGKQAFARVKLRLL---TGTRVEKTEKSTDSAEAGADVVDNML 72
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 351 DRRPFYRTLSDPFVVRGGFGNPHYVLGLRGVAFQGTGTNHTTFRNSGTIDSLDEIPQ- 409
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 73 TYLYNDGEFW-----HFMMNETFEQLSADAKAIGDNAKWLLDQAECIVTLMNGQPISVTPP 128
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 410 ---DNSGAPWMDYSHVLNHTFVRWPG- --ISGSDSW-----RAPMFSWTHR--SATPT 456
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 129 NFVELEIVDTDPGLKGDGTAGTGKRPATLSTGAVVKVPLFVQIGEVK 175
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 457 NTIDPERITQIPLVKAHTLQSG-----TTVVVRGPGFTG-GDILR 494
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: January 12, 2001, 14:59:00
Job time: 64 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 15:01:44 ; Search time 16.19 Seconds
(without alignments)
375.002 Million cell updates/sec

Title: US-09-322-732-2

Perfect score: 975
Sequence: 1 MATYNSNDRFAGLKLMDGE.....QIGEVIKVDRSGEYVSRVK 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	970	99.5	187	1 EFP_ECOLI	P33398 escherichia
2	744	76.3	187	1 EFP_HAEIN	P43771 haemophilus
3	638.5	65.5	187	1 EFP_BUCMP	O51834 buchneria ap
4	455.5	46.7	190	1 EFP2_CHLTPR	O84757 chlamydia t
5	449.5	46.1	192	1 EFP_AQUAE	O67376 aquifex aeo
6	445.5	45.7	190	1 EFP2_CHLNP	Q95711 chlamydia p
7	426	43.7	185	1 EFP_SNP7	Q54760 synecococc
8	424	43.5	187	1 EFP_SNP3	Q55119 synecocyst
9	420	43.1	187	1 EFP_HELPY	P56004 helicobacte
10	417	42.8	187	1 EFP_HELPY	Q92mq5 helicobacte
11	392	40.2	187	1 EFP_MXCTU	P95019 mycobacteri
12	380	39.0	185	1 EFP_BACSU	P49778 bacillus su
13	377	38.7	185	1 EFP_ANASP	Q44247 anabaena sp
14	372	38.2	187	1 EFP_BRELA	Q45288 brevivacter
15	350	35.9	185	1 EFP_TREMA	Q9X284 thermotoga
16	337	34.6	187	1 EFP_TREPA	O83537 treponema p
17	300	30.8	185	1 EFP_BACFR	P70889 bacteroides
18	280.5	28.8	192	1 EFP_BORBU	O51232 borrelia bu
19	252.5	25.9	188	1 EFP_RICPR	Q9Zdt7 rickettsia
20	231	23.7	185	1 EFP1_CHLTPR	O84124 chlamydia t
21	226	23.2	185	1 EFP1_CHLNP	Q9Z900 chlamydia p
22	224	23.0	275	1 YEIP_ECOLI	P33028 escherichia
23	172.5	17.7	190	1 EFP_MYCGE	P47272 mycoplasma
24	170.5	17.5	190	1 EFP_MYCPN	P75085 mycoplasma
25	85.5	8.8	700	1 CAN2_MOUSE	O08529 mus musculu
26	83	8.5	683	1 AMOH_ARTGO	Q59118 arthrobacte
27	81	8.3	1197	1 NIFJ_ANASP	Q06879 anabaena sp
28	80.5	8.3	700	1 CAN2_RAT	Q07009 rattus norv
29	80	8.2	257	1 DNAB_MYCLE	P46394 mycobacteri
30	79	8.1	500	1 ZIPP_DROME	P10379 drosophila
31	78	8.0	1174	1 CLFA_BACTA	Q03746 bacillus th
32	77.5	7.9	376	1 AROC_YEAST	P28777 saccharomyc
33	76.5	7.8	513	1 YA55_SCHPO	Q09735 schizosacch

34	76	7.8	279	1 RL2_LEPIN	Q9xd33 leptospira
35	76	7.8	440	1 OMGP_MOUSE	Q63912 mus musculu
36	76	7.8	531	1 TRPC_PHYPR	P24920 phytophthor
37	76	7.8	608	1 GLMS_ECOLI	P17169 e glucosami
38	75	7.7	642	1 FIMB_YEAST	P32599 saccharomyc
39	74.5	7.6	454	1 THDF_ECOLI	P25522 escherichia
40	74	7.6	363	1 RF2_HELPJ	Q9zmr1 helicobacte
41	74	7.6	874	1 DNAB_MYCTU	P71715 mycobacteri
42	73.5	7.5	609	1 GLMS_HAEIN	P44708 h glucosami
43	73	7.5	382	1 DHGY_CUCSA	P13443 cucumis sat
44	73	7.5	638	1 SYC_HUMAN	P49589 homo sapien
45	72.5	7.4	299	1 Y175_HELPJ	Q9zmq7 helicobacte

ALIGNMENTS

RESULT 1
EFP_ECOLI
ID EFP_ECOLI STANDARD; PRT; 187 AA.
AC P33398;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-32; 70-79 AND 85-116.
RX MEDLINE=92066471; PubMed=1956781;
RA Aoki H., Adams S.-L., Chung D.-G., Yaguchi M., Chuang S.-E.,
RA Ganoza M.C.;
RT "Cloning, sequencing and overexpression of the gene for prokaryotic
RT factor EF-P involved in peptide bond synthesis.";
RL Nucleic Acids Res. 19:6215-6220(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [3]
RP SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [4]
RP CHARACTERIZATION.
RC MEDLINE=97338480; PubMed=9195040;
RX Aoki H., Adams S.-L., Turner M.A., Ganoza M.C.;
RT "Molecular characterization of the prokaryotic efp gene product
RT involved in a peptidyltransferase reaction.";
RL Biochimie 79:7-11(1997).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its


```
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 192 AA; 21807 MW; 9B651E676507CAA0 CRC64;

Query Match          46.1%; Score 449.5; DB 1; Length 192;
Best Local Similarity 48.7%; Pred. No. 2.1e-33;
Matches 92; Conservative 31; Mismatches 65; Indels 1; Gaps 1;

QY 1 MATYYS-NDPRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTGRVEKTEKST 59
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1 MATEIDINRQKDFIEHKGEPRVLDYEHVKPGKGQAFARVKLRLLTGTGRVEKTEKST 60
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

QY 60 DSABGADVDMNLTYLYNDGEFHFMMNTEFQLSADAKAIGDNNAKLLDQAEICIVTLWN 119
   || : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 61 DRIPLADFEQVYATYSYNDGENYFMNTQYDMIAVPKEIEEAKFLKEGMEVIVFLYK 120

QY 120 QPISVTPPNFVELEIVDTPGLKGDGTAGTGGKPATLSTGAVVKVPLFVQIGEVIKVDTR 179
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 121 GQPTIGIELPKHVELQVYVETEPAFKGDTPAQGTAKPAKLETGAVIQVPPFVKEGDIKVDTR 180

QY 180 SGEYVSRVK 188
   : || || |||
Db 181 TGSYVERVK 189

RESULT 6
EFP2_CHLPN STANDARD; PRT; 190 AA.
AC Q9Z711;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE ELONGATION FACTOR P 2 (EFP 2).
GN EFP2 OR CPN0895.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: A501670; AAD19033.1; -.
CC DR
CC INTERPRO: IPR001059; -.
CC DR PFAM; PF01132; EFP; 1.
CC DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 190 AA; 21477 MW; 1A245E50BDAA340C CRC64;

Query Match          45.7%; Score 445.5; DB 1; Length 190;
Best Local Similarity 46.3%; Pred. No. 4.8e-33;
Matches 87; Conservative 42; Mismatches 58; Indels 1; Gaps 1;
```

```
QY 1 MATYYSNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTGRVEKTEKST 60
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1 MVRVSTSEPRVGLRIEIDGQPYLILQNDQFVKPGKGQAFNRKVKNFELTGRVIERTKSGE 60

QY 61 SAGGADVDMNLTYLYNDGEFHFMMNTEFQLSADAKAIGDNNAKLLDQAEICIVTLWN 120
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 61 SVETADIVERSMRLLYTDQEGATFMDDETPEQEVEFWEKLEINRWLLEDITVTLVLYNG 120

QY 121 QPISVTPPNFVELEIVDTPGLKGDGTAGTGGKPATLSTGAVVKVPLFVQIGEVIKVDTR 179
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 121 DVVAEPPFPMELISIAETAPGVRGDTASGRVLRPAVTNTGAKIMVPIFDIDEGELVKVDTR 180

QY 180 SGEYVSRV 187
   : || |||
Db 181 TGSYESRV 188

RESULT 7
EFP_SYNPF7 STANDARD; PRT; 185 AA.
AC Q54760;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE ELONGATION FACTOR P (EFP-P).
GN EFP.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
RN [1]
RP SEQUENCE FROM N.A.
RA Phung L.T., Haselkorn R.;
RT "Genes encoding biotin carboxyl carrier protein and elongation
RT factor P from cyanobacterium Synecococcus sp. PCC 7942.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U59235; AAB82025.1; -.
CC DR
CC INTERPRO: IPR001059; -.
CC DR PFAM; PF01132; EFP; 1.
CC DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20368 MW; BB6880A3D96CF06C CRC64;

Query Match          43.7%; Score 426; DB 1; Length 185;
Best Local Similarity 48.6%; Pred. No. 2.5e-31;
Matches 88; Conservative 25; Mismatches 68; Indels 0; Gaps 0;
```

```
QY 6 SDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTGRVEKTEKSTDSAGA 65
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 4 SDFRTGTTIEIDGAVRWVVEFLHVPKPGKSAFVRTKLKNAKTGNVVEKTRAGETVPOA 63

QY 66 DVVDMNLTYLYNDGEFHFMMNTEFQLSADAKAIGDNNAKLLDQAEICIVTLWNQPISV 125
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 64 VLEKSTLQTYKDGDDDFVFMDETYEEGRLTATIGDRVYKLEKMEANVTWNGQVIEV 123
```


Qy	187	VK	188
Db	186	VK	187

RESULT	10
--------	----

```

EFP_HELPJ
ID      EFP_HELPJ      STANDARD;      PRT;      187 AA.
AC      Q92MQ5;
AD      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      ELONGATION FACTOR P (EF-P).
GN      EFP OR HP0177.
OS      Helicobacter pylori J99 (Campylobacter pylori J99).
OC      Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC      Helicobacter.
RN      [1]
RC      SEQUENCE FROM N.A.
RX      MEDLINE=99120557; PubMed=9923682;
RA      Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA      Smith D.R., Noonan B., Guild B.C., Dejonge B.L., Carmel G.,
RA      Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA      Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA      Trust T.J.;
RA      "Genomic sequence comparison of two unrelated isolates of the human
RT      gastric pathogen Helicobacter pylori.";
RT      Nature 397:176-180(1999).
RL      -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC      TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC      70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC      THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC      THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC      (BY SIMILARITY).
CC      -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC      -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC      -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collabora-
CC      tion between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commer-
CC      cial entities requires a license agreement (See http://www.isb-sib.ch/annotation
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AE001454; AAD05735.1; -.
DR      HSSP; P56635; 1BK6.
DR      PROSITE; PS01275; EFP; 1.
DR      Protein biosynthesis; Elongation factor.
SQ      SEQUENCE 187 AA; 20776 MW; 21733EDFB0FAB331 CRC64;

```

Query Match 42.8%; Score 417; DB 1; Length 187;
Best Local Similarity 45.1%; Pred. No. 1.6e-30;
Matches 82; Conservative 36; Mismatches 64; Indels 0; Gaps 0;

[illegible]

RESULT 11
FEP MYCTI

ID	EPF_MYCTU	STANDARD;	PRT;	187 AA.
AC	P95019;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	ELONGATION FACTOR P (EF-P)			
GN	EPF OR RV2534C OR MTCY159.22.			
OS	Mycobacterium tuberculosis.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=H37RV;			
RX	MEDLINE=98295987; PubMed=9634230;			
RA	Cole S.T., Brosch R., Parkhill J.J., Garnier T., Churcher C., Harris D.,			
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,			
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,			
RA	Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,			
RA	Hornsby T., Jorgels K., Krohn A., McLean J., Moule S., Murphy L.,			
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,			
RA	Rutter S., Seeger K., Skellton S., Squares S., Squares R., Sulston J.E.,			
RA	Taylor K., Whitehead S., Barrall B.G.;			
RT	"Deciphering the biology of Mycobacterium tuberculosis from the			
RT	complete genome sequence.";			
RL	Nature 393:537-544(1998).			
CC	-1- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT			
CC	TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED			
CC	70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING			
CC	THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING			
CC	THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE			
CC	(BY SIMILARITY).			
CC	-1- PATHWAY: PROTEIN BIOSYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.			
CC	-1- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR 283863: CAB06174.1: -

DR TUBERCULIST: RV2534C; -

DR INTERPRO; IPR001059; -.

DR PFAM; PF01132; EFP; 1.

DR PROSITE; PS01275; EFP; 1

KW	Protein biosynthesis; Elongation factor;
SC	SEQUENCE 107 28 20402 197

SQ SEQUENCE 18/ AA; 2040/ MW; 58CBCDIDE0424F58 CRC64;

Query Match 40.2%; Score 392; DB 1; Length 187;

Best Local Similarity 42.88; Pred. No. 2.8e-28;

Matches 80; Conservative 37; Mismatches 68; Indels 2; Gaps 1;

QY 1 MATYNSDFRAGIKIMLDGEPYAVEASEFVKPGKGAFARVKLRLLLTGTRVEKTFKSTD 60
||| : || : || : || : : : ||||| || || : || : || : ||| :

D**b** 1 MAT - - TADFKNGLVLVIDGQLWTITEFQHVKPGKGPFAFVRTKLKNVLSGKVVDKTFNAGV 58

[illegible]

QY 6I SAEGADVDMNLTLYLYNDGEFFHFMNNETFEQLSADAKAIGDNAKWLLDQAECIVTLWNG 120

59 KVDATVDRBRDTTYLYRDGSDNEFEMDSQVFOHPIDPAIVGDAAPETIEGMBVQVAVAEHNC 118

00 KVDIAI VDKRDI I I ETIKGSDI VI MDSQD IEQN EFERLEVGDHANE LLEOMFVQVAFHNG TIO

QY 121 QPISVTPPNFVELEIVDTPGLKGDTAGTGGKPATLSTGAVVKVPLEFVOIGEVIKVDTRS 180

[illegible]

Db 119 VPLYIelpVtVELEvThTEpGLQGRSSAGTKPATLQGTaQINvPLfINTGDKLKVDSRD 178

[illegible]

```

RESULT 13
EFP_ANASP      STANDARD;      PRT; 185 AA.
AC Q44247;
AC Q1-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
[1]
RN SEQUENCE FROM N.A.
RA MEDLINE=93352435; PubMed=8102363;
RA Grnicki P., Scappino L.A., Haselkorn R.;
RT *Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
RT sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
RT protein.*;
RT
RL J. Bacteriol. 175:5268-5272(1993).
CC -1- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA. THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -1- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; L14863; AAA74627.1; -
CC INTERPRO; IPR001059; -
CC PRAM; PF01132; EFP; 1.
CC PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20496 MW; F5F32A12CF106182 CRC64;

Query Match      38.7%; Score 377; DB 1; Length 185;
Best Local Similarity 43.1%; Pred. No. 6.le-27;
Matches 78; Conservative 33; Mismatches 70; Indels 0; Gaps 0;

Qy 6 SNDFRAGLKIMLDGEYAVEASEFVKPGKGQAFARVKILRLLTGTTRVEKTEKSDSAEGA 65
   | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 SNDFRPGVSIVLDGSMVRVIDFLHKVCKGSAFVRTTLKNVQSGKVLKTRAGETVPQA 63
   | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 66 DVDMDNLTLYNDGEFHFHMFNMTFEQLSADAKAIGDNKAWLLDQAEICIVTLWNGQPISV 125
   | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 TLEKITMQHTYKGEDEFFVMDMESVEEGRLSAAQIGDRVKYLKKEGMENVIRWGEQVLEV 123
   | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 126 TPPNFVELEIVDTPGKLGKDTAGTGGKPATLSTGAVKVPFLFVQIGEVIKVDTRSGEVVS 185
   | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 ELANSVWLEIVTQDPGVKGDGTATGTRKPAIVETGATVMVPLFISQGERIKIDTRDKYILG 183
   | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 186 R 186
   |
Db 184 R 184

RESULT 14
EFP_BRELA
ID EFP_BRELA      STANDARD;      PRT; 187 AA.
AC Q45288;
AC Q1-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

Db 179 GSVLGRV 185

RESULT 12

EFF_BACSU STANDARD; PRT; 185 AA.

AC P49778;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DE 30-MAY-2000 (Rel. 39, Last annotation update)

DE ELONGATION FACTOR P (EF-P).

GN Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

RC Kobayashi Y., Mizuno M., Masuda S., Hosono S.,

RA Sato T., Takeuchi M.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 139-185 FROM N.A.

RC STRAIN=168 / JH642;

RC Guerout-Fleury A.M., Gonzy-Treblou G., Stragier P.;

RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE (BY SIMILARITY).

CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; D84432; BAA12558.1; -

CC EMBL; U35252; AAA76718.1; -

CC EMBL; Z99116; CABI4376.1; -

CC SUBTILIST; BGL1460; EFP.

CC INTERPRO; IPR001059; -

CC PFAM; PF01132; EFP; 1.

CC PROSITE; PS01275; EFP; 1.

CC Protein biosynthesis; Elongation factor.

CC SEQUENCE 185 AA; 20454 MW; 60C3B9BF90D2DAB8 CRC64;

CC -----

Query Match 39.0%; Score 380; DB 1; Length 185;

Best Local Similarity 43.3%; Pred. No. 3.3e-27;

Matches 78; Conservative 30; Mismatches 72; Indels 0; Gaps 0;

QY 7 NDFRAGLKIMLDGEPYAVEASEFVKPGKGAFAVKRLRLLTGTFRVEKTFKSTDSAGAD 66

||||| : ||| : ||||| ||||| ||||| ||||| : ||||| : ||

Db 5 NDFRTGLTIDVGIMRVDFQHVHKPGKAFAVRSKRLNLTGATQIEKTFRAGEKVAQA 64

QY 67 VVDNMLTYLNDGEFFHFMNNEFTEQIOLSAKAKATGDNKWLILDDAECIVTLWNQPISTV 126

: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

Db 65 IETKTMQYLVANGQHVHFMDSSTSYEQLELSATQIEEELKYLLENNSHHIMYQDETGLIE 124

||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

QY 127 PPNFVLEIEVDTPGLKGDAGTGCKPATLSTGAVVPLFQVIGVTKVDSRGEVYSR 186

||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

Db 125 LPNTVELKVVETEPGINKDTSAGCTKFAKETGLTVVNVFVFNEDGDTLVVNTSGSYVS 184

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 15:00:08 ; Search time 39.67 Seconds
(without alignments)
555.460 Million cell updates/sec

Title: US-09-322-732-2
Perfect score: 975
Sequence: 1 MATYNSNDRFRAGLIKIMLDGE.....QIGEVIKVTRSGEYVSRVK 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_15:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match %	Length	DB ID	Description	
1	803	82.4	161	2	O69281	citrobacter
2	750	76.9	188	2	O9KNS1	vibrio chol
3	638	65.4	193	2	O9PAM3	xyella fas
4	450.5	46.2	190	2	O9PLH1	chlamydia m
5	445.5	45.7	190	2	O9TOB4	chlamydia p
6	438	44.9	189	2	O9PHW3	campylobact
7	414	42.5	185	2	O9K951	campylobact
8	405	41.5	184	10	O9M917	campylobact
9	378.5	38.8	188	2	O9KXQ9	streptomyce
10	358	36.7	185	2	O9RY32	deinococcus
11	317	32.5	136	2	O9S3F5	helicobacte
12	315	32.3	136	2	O9S3F9	helicobacte
13	315	32.3	136	2	O9S3F6	helicobacte
14	315	32.3	136	2	O9R2M0	helicobacte
15	312	32.0	136	2	O9S3F7	helicobacte
16	312	32.0	136	2	O9R3F4	helicobacte
17	311	31.9	136	2	O9S3F8	helicobacte
18	310	31.8	136	2	O9S3G0	helicobacte
19	291.5	29.9	186	2	O9JU02	neisseria m

20	290.5	29.8	186	2	O9JZQ8	neisseria m
21	283	29.0	187	2	O9RNJ9	zymomonas m
22	235.5	24.2	199	2	O9KSP7	vibrio chol
23	226.5	23.2	189	2	O9PBE1	xyella fas
24	226	23.2	185	2	O9JQ76	chlamydia p
25	221	22.7	185	2	O9PKR6	chlamydia m
26	200	20.5	188	2	O9PQJ3	ureaplasma
27	157	16.1	109	2	P82459	streptococc
28	144.5	14.8	212	10	O9STQ4	arabidopsis
29	87	8.9	161	2	O45051	borrelia bu
30	86.5	8.9	700	11	O54843	mus musculus
31	86	8.8	289	2	O46146	chromohalob
32	84.5	8.7	408	2	O9X1R3	thermotoga
33	84	8.6	307	1	O9UXK6	sulfolobus
34	83	8.5	161	2	O51614	borrelia bu
35	83	8.5	383	2	O9RV55	deinococcus
36	83	8.5	1933	1	O9YEF6	aeropyrum p
37	80.5	8.3	997	1	O9UY74	pyrococcus
38	80	8.2	604	2	O53124	mycobacteri
39	79	8.1	245	5	O17439	boltonia vi
40	78.5	8.1	476	5	O16663	caenorhabdi
41	78.5	8.1	549	2	O82975	caenorhabdi
42	78	8.0	210	2	O9KB07	bacillus ha
43	77.5	7.9	1430	5	O97381	artemia sal
44	77	7.9	303	2	O9Z9P7	artemia sal
45	77	7.9	372	2	O9KFG6	artemia sal

ALIGNMENTS

RESULT 1

- O69281 ID O69281 PRELIMINARY: PRT: 161 AA.
- AC O69281:
- DT 01-AUG-1998 (TREMBlrel. 07, Created)
- DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
- DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
- DE ELONGATION FACTOR-P HOMOLOG (FRAGMENT).
- OS Citrobacter freundii.
- OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
- OC Citrobacter.
- OX NCBI_TaxID=546;
- RN [1]
- RP SEQUENCE FROM N.A.
- RC STRAIN=OS60;
- EX MEDLINE=96032747; PubMed=7559452;
- RA Bishop R.E., Penfold S.S., Frost L.S., Holtje J.V., Weiner J.H.;
- RT "Stationary phase expression of a novel Escherichia coli outer membrane lipoprotein and its relationship with mammalian apolipoprotein D. Implications for the origin of lipocalins.";
- RL J. Biol. Chem. 270:23097-23103(1995).
- RN [2]
- RP SEQUENCE FROM N.A.
- RC STRAIN=OS60;
- RA Bishop R.E., Leski B.K., Hodges R.S., Kay C.M., Weiner J.H.;
- RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
- DR EMBL: U21727; AAC46461.1; -
- DR INTERPRO: IPR001059; -
- DR PFAM: PF01132; EFP; 1.
- DR PROSITE: PS01275; EFP; 1.
- KW Elongation factor.
- FT NON_TER 1
- SQ SEQUENCE 161 AA; 17525 MW; D475D5420DF0BEE4 CRC64;

Query Match 82.4%; Score 803; DB 2; Length 161;
Best Local Similarity 94.4%; Pred. No. 1.4e-65;
Matches 152; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 28 EFVKPGKQAFARVKLRLLTCTRVEKFKSTDSAGADVDMNLTYLYNDGEFHFMMN 87
|||||
DB 1 EFVKPGKQAFARVKLRLLTCTRVEKFKSTDSAGADVDMNLTYLYSDGEFHFMMN 60
|||||

```
Qy 88 ETEFQLSADAKAIGNAKWLLDQAEICVTIWNQOPISVTPPNFVELEIVDTPGLKGDTA 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TTFEQLAADERAGVGNKAKWLLDQAEICVTIWNQOPISVTPPNFVELEIVDTPGLKGDTA 120

Qy 148 GTGKGKATLSTGAVVKVPLFLVQIGEVKIVDTRSGEYVSRVK 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 GTGKGKATLSTGAVVKVPLFLVQIGEVKIVDTRSGEYVSRVK 161

RESULT 2
Q9KNS1 ID Q9KNS1 PRELIMINARY; PRT; 188 AA.
AC Q9KNS1;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE ELONGATION FACTOR P.
GN VC2660.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR NI6961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson L.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi L., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004332; AAF95801.1; -.
DR TIGR; VC2660; -.
SQ SEQUENCE 188 AA; 20576 MW; 73537EA64626B9A2 CRC64;
```

```
Query Match 76.9%; Score 750; DB 2; Length 188;
Best Local Similarity 73.9%; Pred. No. 1.2e-60;
Matches 139; Conservative 26; Mismatches 23; Indels 0; Gaps 0;

Qy 1 MATYNSNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTREKTEKSTD 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MATYSTNEFGKGLKIMLDNPECVILENEYVKPGKGQAFNRVIRKRLTGVLEKTEKSGD 60

Qy 61 SAEGADVVDMLTYLYNDGGEFHFMMNNEFTEQLSADAKAIGDNKWLDDQAEICVTIWNQ 120
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TAEADVVDIDLYLYNDGEYHFMMNSTFEQLAADAKAVGENAKWLVNNTCMLTLWNG 120

Qy 121 QPISVTPPNFVELEIVDTPGLKGDGTAGTGKGPATLSTGAVVKVPLFVQIGEVKIVDTRS 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 NPIAVTPPNFVELEIVDTPGVKGDGTQGTGKGPATLSTGAVRVPLFVQIGEVKIVDTRS 180

Qy 181 GEYVSRVK 188
|||||
Db 181 AEYVGRVK 188
```

```
RESULT 3
Q9PAM3 ID Q9PAM3 PRELIMINARY; PRT; 193 AA.
AC Q9PAM3;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE ELONGATION FACTOR P.
GN XF2473.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
```

```
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Lemus E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-157(2000).
DR EMBL; AE004055; AAF85271.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
SQ SEQUENCE 193 AA; 21380 MW; 8D251D1472A37D06 CRC64;
```

```
Query Match 65.4%; Score 638; DB 2; Length 193;
Best Local Similarity 63.3%; Pred. No. 1.8e-50;
Matches 119; Conservative 26; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MATYNSNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTREKTEKSTD 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 6 MATSGMDVNGKMKILYNAPAVITDTYVKPGKGQAFTRVKYRLKSRVQEVTKSTD 65

Qy 61 SAEGADVVDMLTYLYNDGGEFHFMMNNEFTEQLSADAKAIGDNKWLDDQAEICVTIWNQ 120
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 66 TLEAADVVDDMOYLYSDGEYHFMMQOQTEFQVQADKNGMGAEKWLKGEEQCVVTLWNG 125

Qy 121 QPISVTPPNFVELEIVDTPGLKGDGTAGTGKGPATLSTGAVVKVPLFVQIGEVKIVDTRS 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 126 VPIGVQPPNFVELEIVDTPGLRGDTSGGKGKGPATLETGAVRVPLFVQIGEVKIVDTRS 185

Qy 181 GEYVSRVK 188
|||||
Db 186 GEYVSRVK 193
```

```
RESULT 4
Q9PLH1 ID Q9PLH1 PRELIMINARY; PRT; 190 AA.
AC Q9PLH1;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE TRANSLATION ELONGATION FACTOR P.
GN TC0133.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
```

```

[1]
RX SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RA MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Linher K., Weidman W., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RA pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
[2]
RX SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., White O.,
RA Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Bass S., Linher K., Weidman W., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE002280; AAF39011.1; -.
DR TIGR; TC0133; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
DR SEQUENCE 190 AA; 21550 MW; C5920D53944CD9BD CRC64;

Query Match 46.2%; Score 450.5; DB 2; Length 190;
Best Local Similarity 48.4%; Pred. No. 2e-33;
Matches 91; Conservative 36; Mismatches 60; Indels 1; Gaps 1;

QY 1 MATYNSDFRAGLKIMLDGEPYAVEASEFVKPGQAFARVKLRLLTGTRVEKTFKSTD 60
   | : : : | | | : : | | | : : | | | | | | | | | | | | | | | | |
Db 1 MVRVSTSEFRVGLRIEIDGQPVILQNDVFVKPGQAFNRKIKVKNFLTGRVIERKSGE 60

QY 61 SAEGADVDMNLTLYNDGEFHFNNETFEOLSADAKAIGDNKWLDDQAEICVTLWNG 120
   | | | | | : : | | | : : | | | | | | | | | | | | | | | | |
Db 61 SIETADVREQQRLTYDOEGATFMDDETFEQELIFWFKLENIQWLLDVTYTLVRYNG 120

QY 121 QPISVTPNPFVELEIVDTPGLKGDTA-GTGKGPATLSTGAVVKVPLFVQIGEVKVDTR 179
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 DVISVEPPIFMELSIAETAPGVGTASGRVLKPAVTNTGAKIMVPFIPIFEGELVKVDTR 180

QY 180 SGEYVS RV 187
   : | | | |
Db 181 TGSYESRV 188

RESULT 5
QYQB4 PRELIMINARY; PRT; 190 AA.
AC QYQB4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE ELONGATION FACTOR P (TRANSLATION ELONGATION FACTOR P).
GN EFP.2 OR CP0971.
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
[2]
RX SEQUENCE FROM N.A.
RC STRAIN=AR39;

```

```

RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Linher K., Weidman W., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RA pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AP002548; BAA99103.1; -.
DR EMBL; AE002255; AAF38751.1; -.
DR TIGR; CP0971; -.
KW Elongation factor.
SQ SEQUENCE 190 AA; 21477 MW; 1A245E50BDAA340C CRC64;

Query Match 45.7%; Score 445.5; DB 2; Length 190;
Best Local Similarity 46.3%; Pred. No. 5.6e-33;
Matches 87; Conservative 42; Mismatches 58; Indels 1; Gaps 1;

QY 1 MATYNSDFRAGLKIMLDGEPYAVEASEFVKPGQAFARVKLRLLTGTRVEKTFKSTD 60
   | : : : | | | : : | | | : : | | | | | | | | | | | | | | | | |
Db 1 MVRVSTSEFRVGLRIEIDGQPVILQNDVFVKPGQAFNRKIKVKNFLTGRVIERKSGE 60

QY 61 SAEGADVDMNLTLYNDGEFHFNNETFEOLSADAKAIGDNKWLDDQAEICVTLWNG 120
   | | | | | : : | | | : : | | | | | | | | | | | | | | | | |
Db 61 SVETADIVERSMRLTYDOEGATFMDDETFEQEVVFWKLENIQWLLDVTYTLVLYNG 120

QY 121 QPISVTPNPFVELEIVDTPGLKGDTA-GTGKGPATLSTGAVVKVPLFVQIGEVKVDTR 179
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 DVAVEPPIFMELSIAETAPGVGTASGRVLKPAVTNTGAKIMVPFIPIFEGELVKVDTR 180

QY 180 SGEYVS RV 187
   : | | | |
Db 181 TGSYESRV 188

RESULT 6
QYQB4 PRELIMINARY; PRT; 189 AA.
AC QYQB4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE ELONGATION FACTOR P.
GN EFP.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL; AL139075; CAB75187.1; -.
DR INTERPRO; IPR001059; -.
DR INTERPRO; IPR001884; -.
DR PFAM; PF01132; EFP; 1.
DR PFAM; PF01287; eif-5a; 1.
SQ SEQUENCE 189 AA; 21098 MW; C0BAE2CC7D220300 CRC64;

Query Match 44.9%; Score 438; DB 2; Length 189;
Best Local Similarity 45.7%; Pred. No. 2.7e-32;
Matches 85; Conservative 33; Mismatches 68; Indels 0; Gaps 0;

```


Query Match 32.38; Score 315; DB 2; Length 136;

```

Best Local Similarity 45.2%; Pred. No. 2.6e-21;
Matches 61; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 28 EFVPGKGQAFARVKLRLLTGTTRVEKTFKSTDSAGADVDDMNLTYLYNDGEFHFMMN 87
DB 2 QHVPGKGAFAVRKIKFSLDGKVIKTFHAGDKCEEPNLVEKTMQYLYHGDGTQYQFMDI 61

QY 88 ETEFOLSADAKAIGNAKWLLDQACIVTLWNGQIPISVTPNFVELEIVTDPLGKGDGA 147
DB 62 ESYEQIALNDSQVGEASKWLDGMQVQLLHNDKKAISVDVQVVALKIVETAPNFKGDT 121

QY 148 GTGGKPATLSTGAVV 162
DB 122 SASKKPATLETGAVV 136

RESULT 13
Q9S3F6 PRELIMINARY; PRT; 136 AA.
AC Q9S3F6;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VARIOUS STRAINS;
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DB EMBL; AJ239613; CAB37724.1; -
DB EMBL; AJ239614; CAB37725.1; -
DB EMBL; AJ239615; CAB37726.1; -
DB EMBL; AJ239616; CAB37727.1; -
DB EMBL; AJ239617; CAB37728.1; -
DB EMBL; AJ239618; CAB37729.1; -
DB EMBL; AJ239619; CAB37730.1; -
DB EMBL; AJ239620; CAB37731.1; -
DB EMBL; AJ239621; CAB37732.1; -
DB EMBL; AJ239622; CAB37733.1; -
DR INTERPRO: IPR001059; -
DR PFAM: PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15018 MW; FCC35872F28FA29C CRC64;

Query Match 32.3%; Score 315; DB 2; Length 136;
Best Local Similarity 45.2%; Pred. No. 2.6e-21;
Matches 61; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 28 EFVPGKGQAFARVKLRLLTGTTRVEKTFKSTDSAGADVDDMNLTYLYNDGEFHFMMN 87
DB 2 QHVPGKGAFAVRKIKFSLDGKVIKTFHAGDKCEEPNLVEKTMQYLYHGDGTQYQFMDI 61

QY 88 ETEFOLSADAKAIGNAKWLLDQACIVTLWNGQIPISVTPNFVELEIVTDPLGKGDGA 147
DB 62 ESYEQIALNDSQVGEASKWLDGMQVQLLHNDKKAISVDVQVVALKIVETAPNFKGDT 121

QY 148 GTGGKPATLSTGAVV 162
DB 122 SASKKPATLETGAVV 136

RESULT 14
Q9R2M0 PRELIMINARY; PRT; 136 AA.
AC Q9R2M0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N6;
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DB EMBL; AJ239611; CAB37722.1; -
DR INTERPRO: IPR001059; -
DR PFAM: PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15018 MW; FCC35872F28FA29C CRC64;

Query Match 32.3%; Score 315; DB 2; Length 136;
Best Local Similarity 45.2%; Pred. No. 2.6e-21;
Matches 61; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 28 EFVPGKGQAFARVKLRLLTGTTRVEKTFKSTDSAGADVDDMNLTYLYNDGEFHFMMN 87
DB 2 QHVPGKGAFAVRKIKFSLDGKVIKTFHAGDKCEEPNLVEKTMQYLYHGDGTQYQFMDI 61

QY 88 ETEFOLSADAKAIGNAKWLLDQACIVTLWNGQIPISVTPNFVELEIVTDPLGKGDGA 147
DB 62 ESYEQIALNDSQVGEASKWLDGMQVQLLHNDKKAISVDVQVVALKIVETAPNFKGDT 121

QY 148 GTGGKPATLSTGAVV 162
DB 122 SASKKPATLETGAVV 136

RESULT 15
Q9S3F7 PRELIMINARY; PRT; 136 AA.
AC Q9S3F7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N6;
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DB EMBL; AJ239611; CAB37722.1; -
DR INTERPRO: IPR001059; -
DR PFAM: PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15019 MW; FCC35872F287829C CRC64;

Query Match 32.3%; Score 315; DB 2; Length 136;
Best Local Similarity 45.2%; Pred. No. 2.6e-21;
Matches 61; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 28 EFVPGKGQAFARVKLRLLTGTTRVEKTFKSTDSAGADVDDMNLTYLYNDGEFHFMMN 87
DB 2 QHVPGKGAFAVRKIKFSLDGKVIKTFHAGDKCEEPNLVEKTMQYLYHGDGTQYQFMDI 61

QY 88 ETEFOLSADAKAIGNAKWLLDQACIVTLWNGQIPISVTPNFVELEIVTDPLGKGDGA 147
DB 62 ESYEQIALNDSQVGEASKWLDGMQVQLLHNDKKAISVDVQVVALKIVETAPNFKGDT 121

QY 148 GTGGKPATLSTGAVV 162
DB 122 SASKKPATLETGAVV 136

```

FT NON_TER 136 136
SQ SEQUENCE 136 AA; 15018 MW; FCCDB89CF287829C CRC64;

Query Match 32.0%; Score 312; DB 2; Length 136;
Best Local Similarity 45.2%; Pred. No. 4.8e-21;
Matches 61; Conservative 26; Mismatches 48; Indels 0; Gaps 0;
Qy 28 EFVRPGKQAFARVKLRLLTGTREKTFKSTDSAEAGADVDDMNLTVLYINDGEFWHFMMN 87
 : ||||| || |:: | : |||| : | : ||: : ||||: : ||:
Db 2 QHVRPGKGAAPVRAKIKSFLDGKVIKTFHAGDKCEEENLVEKTMQVLYLHGDYQFMDI 61
Qy 88 ETFEQLSADAKAIGDNAKWLDDQAEICVTLWNGOPISVTPPFVELEIVDTPGLKGDTA 147
 |:|::: : |:|::|: : | | : ||| | |:| | | |||:
Db 62 ESYEQIALNDSQVGEASAKWMLDGMQVQLLHNDKKAISVDVPQVVALKIVKTAPNFKGDT 121
Qy 148 GTGGKPATLSTGAVV 162
 ||||| |||||
Db 122 SASKKPATLETGAVV 136

Search completed: January 12, 2001, 15:00:10
Job time: 130 sec

